

32 No. 3

(New Series No. 147)

January 1972

ANALYSIS

Edited by
C. J. F. WILLIAMS

4 Vols

Picturing unreality

'gavagai'

Can the generalization argument be
reinstated?

Utilitarianism and "conjunctive acts": a
reply to Professor Castañeda

The letter

Symptoms

Indifference and voluntariness

Mannison on doing something on purpose
but not intentionally

True propositions: a reply to C. J. F. Williams

Truth and predication

Black and the inductive justification of
induction

M. R. HAIGHT

CHRISTOPHER S. HILL

WILLIAM G. LYCAN *and*
ANDREW OLDENQUIST

FRED WESTPHAL

JONATHAN HARRISON

JERRY S. CLEGG

JAMES B. BRADY

G. A. SPANGLER

CHARLES SAYWARD

O. R. JONES

BREDO C. JOHNSEN

PASIL BLACKWELL · BROAD STREET · OXFORD

Forty pence

PICTURING UNREALITY

By M. R. HAIGHT

(2)

An artist can represent any individual he pleases; which individual is represented in a picture depends on his intention . . . And he can make his intention to represent Socrates clear, by representing Socrates as having the characteristics traditionally associated with him. But in the way he can represent the subject of his picture as snubnosed, he cannot represent him as a particular real or fictitious person. If he could, it could be part of what a picture showed, that what it showed was real or imaginary—which is obvious nonsense.

(W. Charlton: *Aesthetics*, ch. 3.)

1. **I**N a New Yorker cartoon by Steinberg a small girl walks along a pavement. Behind her on a lead walks an enormous bird. The child and the pavement are drawn in solid line. The bird and its lead are in dotted line. It is immediately obvious—to anyone familiar with the conventions of this kind of art—that the child is real and the bird is something she imagines.

The convention is distinct from the picture. And without it the picture could not represent what it does. But the convention that allows Charlton's painter to make his picture represent Socrates is distinct from that picture too: this must be true of *any* convention, tradition, previous work of art, act or avowal of the artist, etc., (I deliberately leave the list open) which can tell us what the artist means his picture to represent. I see no reason why Steinberg's convention should be outlawed, if Charlton's—or anyone else's—is allowed.

The only things a picture cannot show are those it is logically impossible to represent. (Questions of mere physical possibility, like 'Can I engrave the Taj Mahal in detail on the head of this pin?', are beside the point.) No picture, for example, could at once show you and no one else and me and no one else. If it did, 'This represents you' would be both true and false. No one can paint a *trompe-l'oeil* picture of a round square egg—not because the thing itself is geometrically impossible, but because it is so in a way that would require the painting to be so too, human eyesight working as it does. (Trick paintings of other geometrical impossibilities exist—for example, the wooden triangle reproduced on the cover of the Pelican *New Horizons in Psychology*.) And certainly in representing a thing as imaginary there are limits; here pictures parallel language. I can speak of Pegasus as imaginary—i.e. there is nothing in the world of this description. But the description itself must be of Pegasus *as if* he were real. 'A winged horse'—only real things can have wings, or be horses. Similarly Steinberg can

depict (by a dotted line) the absence, in a world where the child and pavement are real, of anything like the bird which the child imagines behind her. But he can only show us what kind of thing she imagines by drawing what it would look like if, in her world, it were real: a bird about ten feet high. To this extent, then, it is drawn as a real bird. But the dotted-*versus*-solid line convention allows us still to distinguish between the representation of the child as real and the bird as imaginary. It would also allow us to distinguish—where appropriate—a real mermaid from a mythical Editor of ANALYSIS whom she imagines. The subjects' status in real life is irrelevant here—although it may have been important in showing what a dotted outline meant when the device was not yet familiar.

'Imaginary' seems then—after all—to work very much as does 'snubnosed'. One might sum up as follows. The distinguishing features available to a picture-maker are visual. If something visible about a picture could serve as a sign that its subject is ϕ —whatever ϕ may be—then ϕ can in theory be represented. It is true that in some *genres* every visible feature the *genre* will allow has been pre-empted to represent the subject's qualities (or even, specifically, its *visual* qualities). Nothing is left that could count as a sign of its ontological status. This is so, I think, for 'realistic' painting generally, and certainly for *trompe-l'oeil*. But pictorial representation is hardly restricted to these.

2. *Representing the subject of one's picture as a particular real or fictitious person.* This is surely what artists do who portray Mr. Heath as Britannia, or My Lady Batrose as Diana Among Her Nymphs. According to Charlton it can't be done, since (i) it involves showing one's subject, in a picture, to be imaginary (or real); and (ii) this is impossible. We have seen that (ii) is false, if this refers to the subject *in the picture*; while if the subject's *actual* status is what is meant, 'imaginary' is no more uninformative than 'snubnosed'. Wherever a picture can be true to life, it can presumably be false. But (i) seems false, in any case.

Both Steinberg and the painter of Lady Batrose as Diana may be said to picture unreality, but in different senses. Steinberg's bird is shown, by the picture, to be unreal. But what the second artist pictures is unreal simply because he illustrates a fiction. Lady Batrose (in real life, the artist's patroness) is not in fact the goddess Diana. In fact there is no Diana. Still, to represent Lady Batrose *as* Diana surely all the artist needs to do is satisfy, in a single pictured individual, both (i) certain conditions acceptable—in the relevant culture—as identifying a picture of Lady Batrose, and (ii) certain others identifying—in the same culture—a picture of Diana. Conditions for the first might be a considerable degree of likeness, plus the artist's intention to represent (by means of this likeness) Lady Batrose. The second could be identified by a

sufficiency of things like a crescent tiara, bow and quiver, Grecian dress or nymphs. These would allow somebody to identify a painting (correctly) as a representation of the goddess, without knowing whether or not it was also a portrait (and therefore picture) of Lady Batrose. The two sets of conditions are surely compatible.

It might be objected that the Diana-signs relate not to a (mythical) individual, but to some dummy-universal: *Diana-like*, or *dressed as Diana*, or *playing the part of Diana*. I think Charlton himself could not consistently say this, since these are precisely the kind of identifying features he says will do, in the case of the individual Socrates. They are the characteristics traditionally associated with Diana. But anyway the objection will not work. A painter could represent Lady Batrose as *like* Diana, simply by painting her as tall, fair, graceful, in simple flowing clothes, with her hair in a Grecian knot. He could represent her as dressed like Diana and playing the part of Diana between curtains on a stage. None of this is to represent her *as* Diana. And if that is what he wants, he can give signs that make it clear. He might represent Lady Batrose doing something only Diana (traditionally) can, like shining down from the moon at night. Nor is he limited to imaginary figures: he could use a similar method to represent Lady Batrose as Cleopatra in her barge on the Nile.

What can't he do, then? He cannot represent Lady Batrose as not being Lady Batrose. Certainly this is not what he is trying to do when he represents her as Diana: for his picture is of a (fictitious) world where Diana is Lady Batrose. (That is what distinguishes 'My Lady Batrose, Represented as Diana Among Her Nymphs' from 'DIANA, by Sir Josiah Batrose, R.A.—model, Lady Batrose'. The latter is not meant to represent Lady Batrose at all: only Diana, looking like her.) For similar reasons, he cannot represent her as both real and unreal, snub-nosed and not snub-nosed, etc. Nor can he be taken to represent his subject as either real or imaginary in a *genre* that does not allow the appropriate distinguishing features: here the issue is left unsettled. And finally, there are special conditions relating not to 'picture' but to 'portrait'. Our artist can paint a portrait of Lady Batrose which is also a picture of Diana; but he cannot paint a portrait of Diana, because she is not real. He can paint a portrait of Cleopatra (modelled on existing contemporary portraits). But a portrait of Lady Batrose, represented as Cleopatra, cannot I think be more than a *picture* of Cleopatra. (If they were very much alike and he modelled his painting equally on both, he might be said to have done a portrait of both at once. But this is to represent them equally, rather than to represent one as the other.) In short if the only way to picture an individual were to take his portrait, pictures of the kind I have been discussing would indeed be impossible. But—once again—pictorial representation is hardly so restricted.

'GAVAGAI'¹

By CHRISTOPHER S. HILL

WHEN Quine tries to bring leading themes of his philosophy down to earth, he usually relies on his 'gavagai' example. I have no quarrel with the leading themes, but I am suspicious of his observations about 'gavagai'. So as to fix ideas, I open with a version of the example, and follow with an account of Quine's views as to how logic fares under translation. Then, working from these summaries, I argue that the example falls short of illustrating the leading themes.

I

Suppose that a linguist has made a certain amount of progress in translating an exotic tongue. We can imagine that he has found the exotic sentence 'Gavagai' to coincide, in stimulus meaning, with the 'Rabbit' of his own people, and that he has been moved thereby to write five English sentences next to 'Gavagai' in his tentative lexicon. The sentences are: 'Rabbit', 'Undetached rabbit part', 'Rabbit stage', 'Rabbithood', and 'Rabbit fusion'. Not that our linguist is an ordinary one. Rather, in pursuing the philosophy of translation, we are concerned with the behaviour of every imaginable linguist, and therefore, with linguists who are influenced by contemporary domestic philosophers. Such philosophers may demur, or venture only a 'Well, perhaps', when prompted by a rabbit-stimulation and one of the four special sentences; but their behaviour suggests that each of the special sentences might someday become responsive to stimulation, and thereby, to the same stimulations as 'Rabbit' itself.

Having entered the five translations of 'Gavagai' in his lexicon, our linguist comes to the question of whether 'gavagai', as a term now, is to be rendered by our term 'rabbit', by 'rabbit part', by 'rabbit stage', by 'rabbithood', or by 'rabbit fusion'. Since their respective occasion sentences are imagined to coincide in stimulus meaning, the question cannot be answered by checking among patterns of stimulation and assent. Can further behavioural data, say correlation between pointing gestures and speech, help to resolve the problem?

Consider, then, how. Point to a rabbit and you have pointed to a stage of a rabbit, to an integral part of a rabbit, to the rabbit fusion and to where rabbithood is manifested. Point to an integral part of a rabbit and you have pointed again to the remaining four sorts of things; and so on around.²

¹ I wish to thank W. V. Quine and Israel Scheffler for commenting on an earlier version of this note.

² W. V. Quine, *Word and Object* (Cambridge, Massachusetts: The M.I.T. Press, 1960)—afterwards 'WO'—pp. 52, sq.

To get at the extension of 'gavagai', therefore, our friend will need to supplement his pointing with questions about sameness, difference, and number. He will need to ask the equivalent of 'Is this the same thing as that?', or of 'Do we have one rabbit-thing here, or are there more than one?'. But, Quine says, such questioning requires of the linguist a command of the native language far beyond anything that stimulus meaning can teach. Stimulus meaning leaves it open, for example, whether a certain exotic term is to be translated by our identity predicate, or by our 'belongs with' instead. For this reason,

when in the native language we try to ask 'Is this *gavagai* the same as that?' we could as well be asking 'Does this *gavagai* belong with that?'. Insofar, the native's assent is no objective evidence for translating 'gavagai' as 'rabbit' rather than 'undetached rabbit part' or 'rabbit stage'.¹

Thus, neither pointing nor questioning gives a unique answer to the linguist's problem. He will of course make a choice, and in making it he will doubtless be guided by some consideration or other. But any such consideration will pertain less to linguistic behaviour of natives than to his own preference for one of the five domestic terms. Quine concludes that translation of 'gavagai' is indeterminate: that there are five different ways, equally compatible with linguistic evidence, of correlating 'gavagai' with a domestic term. And since the five domestic terms in question differ from one another in reference or extension, he adds that the reference of 'gavagai' is inscrutable (OR, p. 35).

II

Turning now to consider Quine's views about translation of logical words, we find that he is optimistic about signs for the truth functions: he urges that there are 'substantial' behavioural criteria for recognizing negation, conjunction, and the rest (OR, p. 104). On the other hand, he despairs of translating quantificational sentences uniquely. His reason is that they 'depend for their truth on the objects . . . of which the component terms are true; and what those objects are is not uniquely determined by stimulus meanings' (WO, p. 61). Since general terms of different languages cannot be uniquely correlated in point of extension, translation of their quantificational contexts cannot be unique.

Quine allows, however, that there is a point to the question of whether a given exotic expression counts as a quantifier. This attitude emerges as he reflects on substitutional quantification:

Behavioral conditions for interpreting a native construction as existential substitutional quantification, then, are readily formulated. We fix on parts of the construction as candidates for the roles of quantifier and variable; then a condition of their fitness is that the natives be disposed

¹ W. V. Quine, *Ontological Relativity and Other Essays* (New York: Columbia University Press, 1969)—afterwards 'OR'—p. 33.

to dissent from a whole quantified sentence when and only when disposed to dissent from each of the sentences obtainable by dropping the quantifier and substituting for the variable. A second condition is that natives be disposed to assent to the whole whenever disposed to assent to one of the sentences obtainable by dropping the quantifier and substituting for the variable. (OR, p. 105, sq.)

Now, there are many different kinds of substitutional quantification: quantification which interacts with singular terms, quantification which interacts with verbs, and even quantification which interacts with punctuation marks. Quine does not say, in the passage cited, whether it is possible to distinguish between such kinds of quantification at the level of assent and dissent. Let us look into the matter, focusing on quantification which interacts with singular terms.

A linguist will have recognized certain large classes of expressions, known as grammatical categories, in pursuing an inductive definition of the class of native sentences. Will he have any reason to correlate one of these categories with our English category of singular terms? One expects that the answer is 'Yes'. Thus, as Strawson has urged, it is to be expected that truth value gaps, sentences which are neither true nor false, will be seen to depend systematically on members of one grammatical category.¹ There is such a category in English grammar: the category of singular terms.

Whether a sentence is true or false depends on the success or failure of the general term; but the failure of the singular term appears to deprive the general term of the chance of either success or failure. (*Ibid.*)

If, therefore, a study of truth value gaps leads our linguist to category *C*, he will have one reason to correlate *C* with our domestic singular terms. (Truth value gaps can be made into grist for the linguist's mill, as Quine points out, by taking refusal to assent or dissent, exotic indifference, as a sign that a gap is at hand.²)

After looking into truth value gaps, our linguist will perhaps begin to study changes in truth value. If he finds that members of a proper subset of *C* turn up frequently in sentences with fluctuating truth value, and that they never turn up in sentences with constant truth value, he will have found new justification for his initial correlation. For in English, sentences with fluctuating truth value normally have parts which occur in sentences with constant truth value, and parts which do not; and parts of the second kind are by and large singular terms. (Assent and dissent will be clues as to truth value.) Moreover, further justification might be found by looking for another category (perhaps it will turn out to be *D*) which interacts with *C* as follows:

¹ P. F. Strawson, 'Singular Terms and Predication', *Words and Objections*, ed. Donald Davidson and Jaakko Hintikka (Dordrecht: D. Reidel Publishing Company, 1969), p. 100.

² Quine, 'Reply to Strawson', *Words and Objections*, pp. 320-321.

sentences can be built from a single member of *D* by adding two members of *C*, but sentences never result from adding two members of *D* to (one occurrence of) a single member of *C*. If he can find a category which exemplifies this pattern, he will have a reason to correlate it with our transitive verbs, and at the same time, he will have found another point of resemblance between our singular terms and members of *C*.

Not that such cues must coalesce immediately: for example, the study of truth value gaps might focus attention equally on *C* and on *H*. But one expects that these cues and others will in the long run lead our friend to a native category which is closer to our category of singular terms than to any other. And when he finds the right category, he will be one step away from substitutional quantifiers which interact with singular terms.¹

In the quoted passage, Quine describes an experiment for detecting quantifiers. Such experiments will sometimes indicate that informants are using objectual, or normal, quantifiers. For if there is at least one existentially quantified sentence which commands assent, while each of its substitution instances tends to provoke dissent, the exotic people are thereby stamped as objectualists, and at the same time, as believers in nameless objects. On the other hand, when informants are found assenting and dissenting as Quine describes, their quantifiers can be viewed either way. Certainly they could be substitutional quantifiers. It is possible, though, that the natives have an ontology which is bigger than their stock of names, but that no predicate of their theory is satisfied by just the nameless objects. Since both hypotheses accord equally well with patterns of assent and dissent, Quine says that they come to the same thing: in this situation, he is unable 'to distinguish objectively between referential quantification and a substitutional counterfeit' (OR, p. 67).

Be that as it may, we should at least credit the linguist with having learned from his quantificational studies. He can now map exotic sentences uniquely onto logical congruence classes of English sentences. (Sentences are congruent when they have the same logical form up to permutation of substitutional and objectual quantifiers.) This means that he can follow proofs in the exotic language, for his congruence-mapping of sentences determines a congruence-mapping of proofs. In a phrase, he will be able to recognize *logical similarities* between sentences and proofs of the two languages. Linguists can sometimes do better, but they can always do at least this well.

¹ It should be mentioned that the line of thought ending here departs from the letter of Quine's writing on comparative grammar; see e.g. WO, middle of p. 53. It is clear, however, that the above procedure is fully Quinean in spirit.

III

Consider now the primary laws of identity:

- (1) $x = x$
- (2) $[x = y \ \& \ F(x)] \rightarrow F(y)$;

and let us agree to view (2) as a schematic picture of the infinite set of sentences obtainable by replacing 'F' with an English predicate. As is well known, these laws provide for all the general truths about identity, in the sense that each such truth can be proved, in elementary logic, from (1) and one or more instances of (2). But more: they provide a criterion for determining whether or not an exotic predicate should be translated by our identity sign. For it is evident that an exotic relation sign will count as expressing identity if and only if it satisfies these two conditions: First, informants must be disposed to assent to every sentence formed by inserting two occurrences of a singular term in its argument places. And second, when an informant is found assenting to a sentence which contains the sign and two other terms, say 'a' and 'b', and also to a sentence which contains 'a' (or 'b') at one or more places, he must be disposed to assent to a third sentence which is like the second except for having occurrences of 'b' in place of one or more occurrences of 'a' (or of 'a' in place of occurrences of 'b'). Except that it is couched in the language of 'term' and 'relation sign', this two-part criterion might be found anywhere in Quine. But as we saw earlier, this kind of language can be given a behavioural interpretation. It seems, then, that essentially Quinean techniques can be used to determine whether an exotic people share our concept of identity.

When, in Section I, our linguist tried to ask an informant whether this *gavagai* is the same as that one, it seemed that he could just as well be asking whether this *gavagai* part (or this *gavagai* stage) belongs with that part (or with that stage). In view of the criterion, however, we see now that this problem will yield to research. In particular, if the linguist is worried about the first possibility, that he might have asked whether this *gavagai* part belongs with that one, he will presumably proceed as follows. First, he will create a variegated animal, by applying a little dye to the original rabbit. Second, he will consult his lexicon of observation sentences, looking for the correct translation of 'Red'. We can suppose that it turns out to be 'Fleep'. Next, running through his list of exotic predicates, he will pick out the one which is closest in form to 'Fleep'. At this point, he will have three predicates: 'fleep', 'gavagai', and the exotic relation sign, say 'varx', which is troubling him. Blending them with the exotic existential quantifier and a few other logical items (which, we can suppose, turn out to have the same form as corresponding devices of his own people), he will obtain a sentence which looks like (3):

- (3) $(Ex) (Ey) (gavagai(x) \ \& \ gavagai(y) \ \& \ fleep(x) \ \& \ \neg fleep(y) \ \& \ varx(x, y))$.

(3), it will be noted, is similar to (4) and (5) in logical form:

- (4) $(Ex) (Ey) (\text{rabbit}(x) \ \& \ \text{rabbit}(y) \ \& \ \text{red}(x) \ \& \ \neg \text{red}(y) \ \& \ x = y)$;
 (5) $(Ex) (Ey)(\text{rabbit part}(x) \ \& \ \text{rabbit part}(y) \ \& \ \text{red}(x) \ \& \ \neg \text{red}(y) \ \& \ \text{belongs with}(x, y))$.

If *gavagais* are in fact parts of rabbits, (3) will command the assent of anyone stimulated by the variegated animal. (4) is demonstrably false, given an appropriate instance of (2). Hence, if an informant is found assenting to (3), and there is no way of changing his mind by an argument from (3) and an exotic instance of (2), (4) cannot be the correct translation of (3). If *gavagais* are parts of rabbits, therefore, the linguist can with a little effort come to see that his original question, in I, had no bearing on rabbits or identity.

One can feel, initially, that assent to (3) does not rule out all hope of seeing informants as partial to rabbits and of pairing the mysterious 'varx' with identity. For, one thinks, the two occurrences of 'fleep' might have different meanings: 'fleep' might be an indicator word, roughly equivalent to 'red here'. Or again, 'fleep' might mean 'red at one place' at its first occurrence, and 'red at another place' at its second occurrence. Notice, however, that these hypotheses can be tested. If the informant makes no sign that 'fleep' is an indicator word, if no pointing behaviour can be detected, then 'red here' runs afoul of the first principle of translation: translate so as to preserve behavioural patterns. Or take the second suggestion. 'x is red at one place' is substantially a relational expression, and as such, it supports two existential inferences. The suggestion can be tested, then, by checking whether 'fleep(x)' has two existential consequences of its own. If not, then assent to (3) counts as the final word on the matter.

There are analogous procedures to help linguists who wonder whether they have been asking about rabbits or about stages thereof. Still other procedures can resolve indecision as to whether a linguist has been asking about rabbits or about rabbit parts, where the parts in question are spatio-temporal fragments of a big fusion embracing all rabbits. And again, indecision between 'rabbit' and 'place where rabbithood is manifested' can be resolved along similar lines: the rabbit moves; a predicate is found which applies to its first position but not to the second; and an appropriate variant of (3) is presented to the informant.

Further kinds of indecision, as between 'rabbit' and some other term, are imaginable; and they call for slightly different techniques. Thus it

is imaginable that 'gavagai' is a singular term meaning 'the rabbit fusion', and that the linguist has been asking, in effect, 'Is the rabbit fusion here the same as the fusion over there?'. The simplest approach here is of course to check whether 'gavagai' belongs to the category of singular terms. But if additional, independent evidence is wanted, it can be obtained by the following procedure: Point to a red rabbit and a white one, asking 'Fleep?' each time. Any friend of the rabbit fusion can be expected to assent twice; for, holding that both acts of ostension pertain to one and the same thing, a friend of the fusion is bound to hold also that both acts pertain to a *fleep*-ish thing. To spare such people the imputation of craziness, therefore, one would have to render their 'fleep' by 'is partly red', and their 'gavagai' by 'the rabbit fusion'.

It is also imaginable that 'gavagai' is a singular term meaning 'rabbithood', and that the linguist has been asking whether the rabbithood over here is the same as the rabbithood over there (*i.e.*, whether there are two rabbit universals or only one). Insofar as this conceit leads only to indecision between 'rabbit' and 'rabbithood', the issue can be cleared up in about the same way as in the preceding case: essentially the same pattern of assent can be elicited. The only difference is that in translating 'fleep', it will be necessary to resort to a barbarism like 'is partially coinstantiated with redness'.

Counterfeit identity predicates can be distinguished from the real thing, and identity judgements which pertain to rabbits can be distinguished from judgements which pertain rather to the rabbit fusion or to rabbithood. Thus, given that the exotic people are partial to rabbits, their partiality can be recognized. Moreover, deviant partialities can be distinguished from one another. Four deviant general terms have been imagined: one applies to spatial parts of rabbits, another applies to temporal stages, a third applies to spatio-temporal sections of the rabbit fusion, and a fourth applies to places where rabbithood is manifested. Such terms can be distinguished by experiments involving identity predicates. Someone partial to rabbit parts would acquiesce in identity judgements which friends of stages, sections, and places would reject; for he would hold that any twice-ostended part was one and the same object, whereas the others would see a different object each time. (Strictly speaking, a friend of places would agree with a friend of parts until the rabbit moved.) Again, someone partial to stages would acquiesce in identity judgements (involving simultaneous ostensions) that a friend of sections would reject, and would reject judgements that friends of parts and places would accept. A friend of sections would be a pretty negative fellow: he would reject all identity judgements except ones involving simultaneous ostensions directed at the same point. And so on.

We have also imagined exotic singular terms standing for rabbithood

and for the rabbit fusion. By any standard account of these entities, there are parts of the rabbit fusion (*e.g.* the part consisting of your pet rabbit and the tail of mine), which do not instantiate rabbithood. Once one such part is specified by a series of pointing gestures and identity statements, a couple of questions about it will relieve indecision between 'rabbithood' and 'the rabbit fusion'. And by the same token, the answers will indicate whether informants are working with a relation sign for pairs of instances and universals, or with one for pairs of parts and fusions. (Should there be non-standard views about rabbithood or the rabbit fusion in the background, the problem would call for more subtle techniques.)

When certain natural assumptions are satisfied, therefore, patient scrutiny can put our linguist in touch with the extension or reference of 'gavagai'. The major assumptions are: that there is an exotic grammatical category similar to our category of singular terms; that the exotic vocabulary is sufficiently rich to support a full-fledged identity predicate; that exotic pointing gestures can be identified and correlated uniquely with our own. Insofar as exceptions and borderline cases are imaginable, there is still room for a Quinean claim: not, perhaps, for the claim that reference to rabbity entities is inscrutable, but certainly for the claim that such reference is local to a proper subset of the range of conceptual schemes (*cf.* WO, p. 53). On the other hand, one would have to travel in thought to Mars or beneath the sea to find exceptions; by and large, reference to rabbity entities is common coin.

Quine's leading themes pertain less to rabbits than to the nature of scientific theories; and as is well known, they do not imply his observations about 'gavagai'. Yet the example tends in its salience to divert attention from the themes. I hope that this note will help to clear the air.

Harvard University

CAN THE GENERALIZATION ARGUMENT BE REINSTATED?

By WILLIAM G. LYCAN and ANDREW OLDENQUIST

CONSIDER these three principles:

(GC) If the consequences of everyone's doing $[A]$ would be undesirable, then not everyone has the right to do $[A]$.

(GP) What is right for one person must be right for any . . . person . . .

(GA) If the consequences of everyone's doing $[A]$ would be undesirable, then no one has the right to do $[A]$

Marcus Singer claims (in *Generalization in Ethics* (New York: Alfred A. Knopf, 1961), ch. IV) that (GC) and (GP) are true¹ and that (GA) is entailed by the two taken together. These claims have been criticized in the literature² but never, so far as we know, subsequently defended. Let us now test them ourselves by translating the three principles into a simple quantified deontic notation:

$Up =_{df}$ The situation (its being the case that) p would have undesirable consequences.

$Op =_{df}$ It is morally obligatory that $p =$ It ought to be the case that p .

$Ax =_{df}$ x does action A . (Variables range over persons.)

Now (GP) is just the familiar Universalizability Principle; and Singer says (p. 66) that, in the antecedents of (GC) and (GA), 'everyone's doing $[A]$ ' is to be taken collectively rather than distributively. So we may express (GP) and (GA) as follows:

(GP*) $(\exists x)O-Ax \rightarrow (y)O-Ay$

(GA*) $U(x)Ax \rightarrow (x)O-Ax$

(GC), owing to the modal ambiguity of the terms expressing its consequent, has two possible readings, even after Singer's own explanations have been taken into account.

(GC*-1) $U(x)Ax \rightarrow O-(y)Ay$

(GC*-2) $U(x)Ax \rightarrow -(y)O-Ay (= U(x)Ax \rightarrow (\exists y)O-Ay)$

¹Actually, Singer adds, as he must, a *ceteris paribus* qualification to (GP) and a corresponding one to (GA). But, since these do not figure essentially in the arguments we shall be discussing, we shall ignore them for the sake of simplicity.

² Gertrude Ezorsky, review of *Generalization in Ethics*, *Journal of Philosophy*, Vol. LX (1963); A. P. Griffiths, review of *Generalization in Ethics*, *Philosophical Books*, Vol. III (1962), and 'The Generalization Argument', *ANALYSIS*, Vol. XXIII (1962-1963); Robert L. Holmes, 'On Generalization', *Journal of Philosophy*, Vol. LX (1963); David Keyt, 'Singer's Generalization Argument', *Philosophical Review*, Vol. LXXII (1963); George Nakhnikian, 'Generalization in Ethics', *Review of Metaphysics*, Vol. XVII (1963-1964).

It may be argued that the "transcendental deduction" of (GA) (or (GA*)) is unacceptable (though not necessarily invalid, as some commentators on Singer have thought) on each reading of (GC). Let us look first at (GC*-1). Plainly Singer believes it is true, since it is a trivial consequence of a more general principle that he evidently accepts: $Up \rightarrow O-p$. We hold this principle too, pending the resolution of problems in interpreting 'undesirable'. Now, does (GA*) follow logically from (GC*-1) and (GP*)? Singer appeals to the contrapositive of (GP*):

$$\begin{array}{l} (GC^*-1) \quad U(x)Ax \rightarrow O(y)Ay \\ (C-GP^*) \quad -(y)O-Ay \rightarrow (z)O-Az \text{ (contrapositive of (GP*))} \\ \hline \therefore (GA^*) \quad U(x)Ax \rightarrow (z)O-Az \end{array}$$

This inference is not obviously valid. It could be valid if the following conditional is analytic:

$$(CD) \quad O(\exists y)-Ay \rightarrow (\exists y)O-Ay,$$

which says that if there ought to be someone *or other* who refrains from doing A, then there is some particular person who is forbidden to do A¹. (CD) is certainly not plainly true. A putative counter-example can be constructed on the basis of a case put forward by A. P. Griffiths (*op. cit.*).

Before examining it, however, let us try to prove (GA*) by substituting (GC*-2) for (GC*-1) in the preceding argument. The result, it turns out, is trivially valid:

$$\begin{array}{l} (GC^*-2) \quad U(x)Ax \rightarrow -(y)O-Ay \\ (C-GP^*) \quad -(y)O-Ay \rightarrow (z)O-Az \\ \hline \therefore (GA^*) \quad U(x)Ax \rightarrow (z)O-Az \end{array}$$

But we may question the truth of (GC*-2). It says that if the consequences of everyone's doing A would be undesirable, then there is some particular person who is forbidden to do A. This is so far no more obvious than (CD), though it is defended cautiously by David Braybrooke²; for it, too, runs into Griffiths' purported counter-example.

The case is set up by Griffiths as follows: Let our universe contain just three persons, J, K and L. They have a "two-seater", by which is meant a car that only two individuals can feasibly fit into (let us suppose, to amplify on Griffiths a little, that, were all three individuals to get into the car simultaneously, disaster would ensue, in the form of broken springs and windows). Now $(\exists y)O-Ay$

¹ What (CD) does, in effect, is merely allow us to deduce (GC*-2) from (GC*-1) and thereby make the subsequent argument valid (see below).

² 'Collective and Distributive Generalization in Ethics', *ANALYSIS*, Vol. XXIII (1962-1963).

may be thought to follow [from $U(x)Ax$] because, if the undesirable consequences are to be avoided, then either J , or K , or L , must refrain from doing $[A]$; and if . . . one ought to avert disastrous consequences, then either J , or K , or L *ought* not to do $[A]$. But this is a mistake. For J can reason quite validly as follows: I may do $[A]$; because L will not do it, and hence if I do no disastrous consequences will follow. L can reason similarly in relation to K , and K in relation to J . This works because ' x may do $[A]$. . . but he will not' is not any kind of contradiction. Nor then is the conjunction 'If J, K and L all do $[A]$, there will be undesirable consequences; one ought to avoid undesirable consequences; but not everyone will do $[A]$ in fact, therefore it is not the case that J , or K , or L ought to refrain from doing $[A]$, in order to avoid consequences which will not ensue.' (pp. 114–115)

The point, we take it, is this: In Griffiths' hypothetical case $U(x)Ax$ is true; but there does not seem to be any particular individual (J, K , or L) in the domain who satisfies ' $O-Ay$ ', since (other things being equal) neither J, K nor L is specifically forbidden to get into the car. Surely no one of them would blame another for getting into it alone in order to go out for beer. If this is right, (GC^*-2) is false, and therefore our second argument for (GA^*) is unsound.

Although Griffiths does not mention it himself, we believe that if his case is an effective counter example to (GC^*-2) , it also counts against (CD) . In the situation envisaged, $O(y)Ay (= O(\exists y)-Ay)$ is presumably true, since it can be derived from $U(x)Ax$ and $Up \rightarrow O-p$. But, for the reasons Griffiths gives, it may appear that $(\exists y)O-Ay$ is false in our three-man universe. If so, (CD) is false, our (GC^*-1) argument for (GA^*) is also unsound, and Singer has neither argument left.¹

Before seriously criticizing Griffiths' position, let us get two common but spurious objections to it out of the way. First, of course, anyone who *already holds* (GA) would deny the possibility of Griffiths' case outright, since (GA) rules that *each* of our three individuals is specifically forbidden to get into the car². But to presume (GA) in meeting a putative counter-example to a step in the only argument for (GA) would be viciously circular. A second natural response to Griffiths' case might be, ' J , for example, may well be specifically forbidden to get into the car if K and L are already sitting in it'. Certainly we may admit that, *under certain*

¹ There is actually another way of making Singer's deduction obviously valid (suggested by George Nakhnikian, *op. cit.*, pp. 448–449): render (GP) as $(GP^*-2) O(\exists y)-Ay \rightarrow (\exists x)O-Ax$. However, (i) it is hard to see how (GP^*-2) could possibly be regarded as an adequate paraphrase of (GP) , and (ii) there is very little antecedent plausibility to (GP^*-2) unless we rely on (CD) . Nakhnikian argues cogently against (GP^*-2) in Part III of his paper.

² We are supposing that the case of the two-seater is non-invertible, i.e., that, were none of the three individuals to get into the car, no disaster would ensue.

We need not stipulate, as Griffiths seems to, that J must *know* that K and L are uninterested in using the car. (That would count as a "special justifying reason", in Singer's *ceteris paribus* terminology, and thereby automatically save Singer from the counter-example.) We need stipulate only that K and L are not sitting in the car.

circumstances (when other things are *not* equal in certain respects), *J* might be forbidden to get into the car. But this is just to say that $(\exists y)O-Ay$ follows from the conjunction of $O-(x).Ax$ with certain other premises; and that, though true, is unhelpful.

Let us move on, then, to a third objection to Griffiths which will prove much more telling. The point of offering (GA) as a normative theory or as the cornerstone of one is to accommodate certain intuitions which are not caught by theories that ignore the question of what would happen if everyone were to do *A*. In particular, (GA) is offered as a more intuitive alternative to traditional utilitarianism. Singer obviously respects "undesirable consequences" as having moral importance, as the utilitarian does; but he believes, probably correctly, that most of us feel (as a matter of justice) that one does wrong to evade paying one's taxes, pollute the air, cheat on exams, etc. *whether or not* some individual instance of such an action has undesirable consequences (see pp. 207ff). So Singer holds that a normative theory is defective if it makes the moral worth of an act a simple function of the desirability of the act's causal consequences.

Now why does Griffiths' proffered counter-example seem plausible? Why do we want to say that *J*, for instance, has no obligation to stay out of the two-seater? Presumably because his getting in would have no bad consequences and perhaps beneficial ones (unless, of course, we add the extra premise that *K* and *L* are already in the car). But one of Singer's normative theses is that individual actions can be wrong in the absence of undesirable consequences. He does not claim this simply on the basis of (GA), which move would be circular; he maintains that it is one of our common moral intuitions which motivate rejection of naive utilitarianism in favour of (GA) in the first place. So it would seem that Griffiths' "counter-example" tacitly begs the question, since it relies on an intuition which Singer has *already* claimed to be at odds with a more basic and reliable one. Our objection could be put in this way as well: Griffiths' "counter-example" is a telling one *if* act-utilitarianism is true, or if some other theory is correct which implies that an individual action is wrong only if it has undesirable consequences. But it does not support such a theory, unless one first accepts Griffiths' invitation to share his intuition, one which Singer and many others already reject.

This shows neither that act-utilitarianism is false, nor that (GA) is true, nor yet that Singer's derivation of (GA) is any good. And some philosophers are temperamentally suspicious of charges of "tacit question-begging", since, when an issue has become as involved as this one has, the question of who claimed what first is pedantic. Still, the objection to Griffiths does prove something important: that the (GA) issue boils down so far to an unresolved conflict of intuitions

which must be addressed afresh. The matter is thus not dead, as some commentators have thought, and it rests heavily (we hope we have shown) on an older and less sophisticated controversy.

We shall close by tying the conflict of intuitions in with a few others which have to do with (CD). The first concerns the appropriateness of our deontic notation itself. The obligation operator 'O' operates on sentences standing for states of affairs—which indicates that states of affairs or facts, as well as the occurrences of actions, are the items which admit of moral evaluation. Two criticisms might be made: First, it might be objected that only individual persons can be morally obliged and only individual actions can be morally right or wrong, and therefore that the operator 'O' cannot intelligibly be applied to a sentence which stands simply for a state of affairs. Hence, sentences of the form, " $O(\exists x)p$ " are unintelligible, and so (GC*-1) and (CD) are unintelligible.

If this objection is sound it counts against the (GC*-1) argument for (GA), but it does not count against the (GC*-2) argument, since the latter contains no occurrence of 'O' whose scope is a closed sentence.

In reply to the objection, we offer an example which, though it questions the truth of (CD), supports its intelligibility. (CD) says that if it ought to be the case that there is someone who refrains from doing A , then there is someone who ought to refrain from doing A . Our apparent counter-example is based on the "ought" implies "can" principle: Suppose it is the case that the war is immoral and ought not to be permitted to continue. Presumably, then, there ought to be someone (or more than one person) who does not permit it to continue, i.e., who stops it. Suppose also that, though it is causally possible that there be someone who stops the war, in fact there is no person who is causally able to stop the war. (We assume the falsity of $Mc(\exists x)Ax \rightarrow (\exists x)McAx$, where 'Mc' is the causal-possibility operator. Here is a counter-example: There happens to be no one now in the world who is causally able to lift 500 pounds; but it is causally possible that such a person be born, and thus causally possible that such a person exist.) It follows from this last assumption (by "ought" implies "can") that there is no person who is obliged to stop the war¹. This case makes $O(\exists x)Ax$ true but $(\exists x)O-Ax$ false (where ' A ' stands for '... permits the war to continue'), thus blocking the kind of exportation codified by (CD).

As usual, our "counter-example" rests on an intuition (that there ought to be someone who stops the war, even though there is no actual person who has the ability to do so) which surely will be challenged. But it is one that must be taken seriously; it goes without saying that those who share this intuition would hold that $O(\exists x)Ax$ and the like are *intelligible*. And once again, we suggest that more work in

¹ We owe this particular example to our colleague, Charles Kielkopf, from whose conversation we have profited.

moral theory is required before one can adjudicate between the conflicting moral feelings.

Parenthetically, we can offer one more shot at (CD), by invoking the notion of responsibility. $(\exists y)O-Ay$ surely involves individual responsibility, in that it entails that someone in particular is responsible for refraining from doing A . Upon being told, 'There's someone who is forbidden to do A ', it would not be out of place for us to ask, 'Which of us is it that bears the burden?' But $O(\exists y)-Ay$ does not seem to entail anything about personal responsibility. If we were told, 'There ought to be someone in the world who doesn't do A (if only God would create a man with such will-power)', a reply of 'Which of us is it?' would be shockingly odd. The partisan of (CD) might rejoin that the oddity vanishes as soon as we see that $O(\exists y)-Ay$ entails $(\exists y)O-Ay$, but it seems trenchant to us, (Naturally, it does not count at all against the (GC*-2) argument.)¹

The second objection to our notation is similar to the intelligibility objection but less drastic: that our obligation operator, though it forms intelligible (and sometimes true) sentences, marks a different sense of obligation from that which is ordinarily applied to persons or acts. It is a different sense of obligation because, unlike the ordinary sense, it applies to states of affairs and to occurrences of actions. It has little to do with persons and their mundane responsibilities; it is rather an abstract matter of the evaluative choice between two or more fancied possible worlds (cf. the sense of 'ought' that G. E. Moore had in mind when he said that what is intrinsically good ought to exist and that an ugly universe ought not to exist).

If one concedes this point, one might then choose not to allow 'O' to operate on closed sentences, in order to guard against the alleged equivocation. This would be tantamount to giving up the (GC*-1) argument and relying wholly on the (GC*-2) argument. On the other hand, we are hard put to say either how the objection could be proved or how it could be refuted, since there is at present no general way of settling charges of ambiguity. We suspect that any decision on this particular issue concerning the ambiguity of 'ought' would rest squarely on prior moral intuitions—once more.

We have tried to show that the dispute to which the (GA) issue boils down cannot be resolved without a fresh start in normative theory. Singer's deduction of (GA) begs some important normative and meta-ethical questions, but so do the criticisms of it which we have discussed. The dispute, if it has been buried in the literature, has been buried alive.

¹A similar distinction is made by I. L. Humberstone in 'Two Sorts of "Ought"s', *ANALYSIS* XXXII, No 1. See also Hector-Neri Castañeda, 'On the Semantics of the Ought-To-Do', *Synthese*, Vol. 21 (1970).

UTILITARIANISM AND "CONJUNCTIVE ACTS": A REPLY TO PROFESSOR CASTAÑEDA

By FRED WESTPHAL

ONE would have thought that Utilitarianism already had more than an ample number of problems, but Professor Castañeda has claimed to have found yet another. I think, however, that the Utilitarian need not lose any sleep over this one.

Professor Castañeda's argument¹ runs as follows:

- (I) For every kind of obligation, it is a principle that 'X is obliged to do both A and B' entails 'X is obliged to do A and X is obliged to do B.'
- (II) The Utilitarian principle states that X is obliged to A in circumstances C, if and only if X's doing A in C will bring about a greater balance of good (pleasure, etc.) over bad (pain, etc.) than his doing any other alternative action open to him in C.
- (III) Supposing it is the case that X is morally obliged to do a conjunctive act P and Q, it follows from principle (I) that X is obliged to do P and X is obliged to do Q.
- (IV) But according to the Utilitarian principle (II above), X is obliged to do P if and only if it is the *optimific* act and he is obliged to do Q if and only if *it* is the *optimific* act.²

Although Castañeda does not state exactly how we are to understand the expression 'conjunctive act', it is reasonably clear from things he has said in another place³ what he has in mind. We would be commanding a "conjunctive act" if we said, for example, 'Open the door and keep it open' and likewise if we said, 'Open the window and shut the door'. It should be noted that there is a significant difference between these two examples, which Castañeda himself has used. It is clear from the meanings of the words involved in the first "conjunctive act" that it would be more accurately stated as follows: 'Open the door and *then* keep it open'. In the case of the second example, one would have to know something of the context in which one was told to open the window and shut the door in order to tell in which sequence if any the two actions were to be done. It could be, of course, that the order in which they are to be done is left to the discretion of the agent involved.

Let us take each of these two examples and fit them into Castañeda's argument which is purported to refute Utilitarianism. Suppose it is the

¹ 'A Problem for Utilitarianism', *ANALYSIS*, Vol. 28, No. 4, pp. 141-42.

² For the sake of convenience, I shall use the expression '*optimific* act' in place of 'the act which will bring about a greater balance of good over bad than any alternative act open to the agent in the circumstances'.

³ 'The Logic of Change, Action, and Norms', *Journal of Philosophy*, Vol. 62, No. 13, 1965, pp. 333-44.

case that a given person X in circumstances C is morally obliged to open the door and keep it open. We may suppose that this person is obliged to open the door of a burning schoolhouse and keep it open so that it doesn't jam and thereby trap the school children inside the blazing inferno. On Castañeda's first principle, the statement 'X is obliged to open the door and keep it open' entails 'X is obliged to open the door and X is obliged to keep the door open'. Now, Castañeda would go on to say, it is impossible for two different actions to be *the* action which produces more good than any other action open to the agent in a given set of circumstances. And following the Utilitarian principle, we must conclude that both of these actions are *the* optimific act because both are obligatory. Utilitarianism, therefore, entails a contradiction and should be regarded as a false moral principle.

The Utilitarian, however, can escape this charge unscathed if he is a little more precise in the way he states X's obligations. The time at which he has each of his obligations is extremely important. Since one cannot *keep* open what has not been already opened, X's obligation to open the door must be carried out or fulfilled first. It is clear that (a) he cannot both open the door and see to it that the door is *kept* open at the same time, and (b) he cannot *keep* the door open before he has opened it. Having made this clear, the Utilitarian should then go on to state X's obligations in the following way: At time T_1 , X is obliged to open the door and at time T_2, T_3, T_n , X is obliged to keep it open. (How long X is obliged to keep the door open will depend, of course, upon how long it takes all of the people inside the schoolhouse to get out). If the Utilitarian specifies the nature of X's obligations in this way, no contradiction results because of his application of the Utilitarian principle. At time T_1 , the optimific act for X to perform is opening the door and at T_2 the optimific act for X is seeing to it that the door stays open. Both actions are optimific and, therefore, obligatory, but they are so at *different times*.

Suppose we consider the second of our two examples of a "conjunctive act", examples which, you will recall, Castañeda has employed in another essay. Suppose X is a teacher at the school which burned down. Prior to the disaster he was told by his superiors that at the beginning of the first class period of each day when there was warm weather, it was his duty to open the window and shut the door. The opened window would foster good ventilation and the closed door would cut down on noise and distractions in the hallway. Both would contribute significantly to the students' learning processes. Perhaps we can assume that a Utilitarian would consider these two actions by the teacher as the best ones he could do in the circumstances, that is, at the beginning of the first class period during warm weather. Nothing *he* could do *in those*

circumstances would do more for his students' education which a utilitarian would consider productive of significant good.

Following Castañeda's line of argument, X is obliged to open the window and he is obliged to close the door. But if he is obliged to do both, then both must be the optimific act. Again, a contradiction seems to result from applying the Utilitarian principle. But just as in the first example, a way out for the Utilitarian is clear. Since under normal circumstances it is physically impossible for one person to open a door and a window at the same time, Mr. X could not be obliged to perform both of these actions at the same time. (If there were some mechanical device which when activated opened the window and closed the door, this would be a different matter.) Again the Utilitarian could simply point out that at time T_1 , Mr. X is obliged either to open the window or to close the door and at T_2 he is obliged to do whichever of the two he did not do first. If it is clear from the context in which these obligations hold that the order in which he does the acts is of no consequence, then his obligation is simply to do one or the other first and follow it up with the one remaining.

It hardly needs pointing out that there is a feature of the Utilitarian principle which it shares with other theories of moral obligation. A necessary condition for a person's being obliged to perform a given act is that that act must be logically and physically possible for the person in question to perform at the time the obligation is supposed to hold. One way for the Utilitarian to avert the attack by Castañeda is simply to deny that a person could be obliged to perform at a given time any "conjunctive act" both of whose conjuncts could not (logically or physically) be done at that time by the same person. But if the Utilitarian agrees to regard X's obligation to perform P and Q as a case of being obliged to perform a "conjunctive act", he must be careful *not* to agree that the only thing this means is that "X is obliged to do P and X is obliged to do Q". If he states precisely *when* X ought to do each of the actions P and Q, he escapes the trap into which Castañeda thinks the Utilitarian must necessarily fall.

At this point an objection could be raised. Suppose it is the case that I have got the Utilitarian off the hook by using examples of one class of "conjunctive acts", namely those which involve two actions which cannot either logically or physically be done at the same time by the same person. There is another type of conjunctive act which hasn't been considered. There are some conjunctive acts whose conjuncts *can* logically and physically be done at the same time by the same person and we can think of some of these as obligatory acts for a given agent on Utilitarian grounds. Suppose X is obliged to turn off valve A and turn on valve B at the same time. Unless both are turned in the right way at the same time, immense pressure will immediately result and cause an

105/An 13

explosion. To say 'X is obliged to turn valve A off and valve B on' entails 'X is obliged to turn valve A off and X is obliged to turn valve B on'. But turning valve A off cannot be the optimific act, which it must be if it is obligatory. And turning valve B on cannot be the optimific act, which it must be if it is obligatory. (Turning only one valve at a time would not prevent an explosion.) Hence the Utilitarian is in the uncomfortable position of saying that a certain act is obligatory but having to admit that it is not optimific. Can the Utilitarian get out of *this one*?

I think that a way of escape is open and it is similar to the one which was suggested for the other type of obligatory conjunctive acts. The obligation to turn valve A off and valve B on should be referred to in such a way that no ambiguity surrounds the word 'and'. The obligation which X has is as follows: At time T_1 X is obliged to turn valve A off and at time T_1 he is obliged to turn valve B on. Since he is obliged to do both acts *at the same time*, doing both *at the same time* must be (a) logically and physically possible for the same person, and (b) the action which produces a greater balance of good over bad than would result from turning each of the valves at different times. There is no difficulty here for the Utilitarian since neither act alone is the optimific one; only by being performed at the same time do each of the conjuncts become part of an obligatory conjunctive act. It is simply misleading to say that 'X is obliged to turn A off and B on' entails 'X is obliged to turn A off and X is obliged to turn B on' and then let it go at that. X's obligation is not *merely* to turn A off and B on but to turn A off *while* turning B on. Since neither act alone is optimific, it follows that neither act alone is obligatory. And since this is the case, the Utilitarian is not involved in saying something contradictory, namely that two different actions considered separately could both be *the* optimific act. He can, however, consistently hold that two actions could in certain circumstances (i.e., when performed at the same time) comprise the optimific thing to do.

University of Miami

THE LETTER

By JONATHAN HARRISON

ONCE upon a time a young man—we shall call him Jack—received a letter through the post. It was not like the other letters he received that morning, which were the usual collection of bills and circulars, and a communication from one of his aunts. It was on faded parchment, with a slight aroma of perfume, written in a copper-plate handwriting, and stamped with a penny stamp with the head of Queen Victoria on it. The date on the postmark was 1st January, 1851. Despite all this, it was addressed, mundanely enough, to *The Occupier*. He opened it, and found in it a letter from a woman, whom we shall call Jill. In it she explained, though in more eloquent language than I have at my command, that she was betrothed to a man she did not love, and had given up all hope of meeting by normal means anyone who would be a meet object of her affections. Nevertheless she did not entirely despair that one day she would find such a man, and recently there had grown up inside her the irresistible feeling that, if only she could communicate with the occupier of the house in which Jack now resided—it was an old house, and even in 1851 was not new—she would find someone with whom she could somehow or other enter into a relationship more satisfying than any other she was likely to know. Hence, on the 1st January, 1851, she had written and posted the letter which, on the 2nd January, 1971, Jack received. It had, of course, been lost in the post, and when eventually discovered had been punctiliously delivered, though with no apology, but a polite request from the postman for an excess postal charge.

Now here the story might have ended had not Jack himself been in a similar lonely position to that of Jill, and had he not inherited from his Celtic ancestors a readiness to believe in possibilities which would seem incredible to less volatile and credulous temperaments. So he replied to her, explaining his position, and expressing the hope that, though the space-time interval between them precluded a more intimate union, they might, nevertheless, enter upon and develop a relationship to the extent that this could be done by correspondence only. Jack posted his letter on the 2nd January, 1971—by a sudden inspiration, in an old post box down the road, which clearly had not been used for many decades.

On the 3rd January, 1971, nothing happened, but on the 4th January, 1971, Jack received a reply from Jill, in which she expressed her compliance with his proposals, and her delight in the prospect of a relationship which, though spiritual, offered her the solace of companionship in the hard and lonely years which she anticipated. This correspondence,

to the mutual satisfaction of both parties, continued until Jill died of a combination of consumption and excessive indulgence in childbirth ten years or so later.

* * *

I tell you this story, not simply because I have a romantic disposition, but because it raises some, though not all, of the problems discussed in *Dr. Who and the Philosophers*.¹ It does not raise the problem of personal identity. The whole of Jack's life is lived out in the twentieth century, and the whole of Jill's in the nineteenth, and no section of the life of either is in any way displaced. It does, however, raise the problem of the identity of the letters written by Jack—though there is no problem about the identity of the letters written by Jill; they just get lost in the post. But Jack's letters, let us suppose, miraculously disappear in the post box in 1971, and they, or ones just like them, miraculously appear in the post bag in 1851. The problem whether the letter Jill receives is the same as the letter Jack posted may perhaps be solved by distinguishing between a type and a token sense of 'letter'. In the token sense of 'letter' the letters Jack posts are not the same as the letters Jill receives, just as, in the token sense of 'telegram' the telegram Jones writes is not the same telegram that is delivered to Smith. But in the type sense of 'letter', the letter Jill receives is the same letter as the one that Jack writes, just as the book I read is the same as the book read by someone in Australia, and the telegram Jones writes is the same as the telegram Smith receives. Now perhaps—and I make the suggestion very tentatively, for it is an extremely bizarre one—if sections of people's history were frequently displaced, as in Professor Williams's example² of the numerous men each of whom came to have the memories, dispositions, abilities and character traits of Guy Fawkes, we would come to distinguish between a type and a token sense of 'person'. Hence, in the token sense of 'person' the man in the Portland Building with Guy Fawkes's memories and the man in the Trent Building with Guy Fawkes's memories are not the same person, but in the type sense of 'person', they are the same person, and whichever one of them we meet, we are meeting Guy Fawkes.

The *second* problem, raised both by my story and by *Dr. Who and the Philosophers*, is the problem of backwards causation. Communication is, in part at least, a causal relationship, so if Jack, in the twentieth century, is to have succeeded in communicating with Jill, in the nineteenth, the things Jack says in his letters must causally affect Jill's thoughts and feelings and the things she says in her replies, even though Jill's replies are written before Jack's letters. We must not suppose that the effect of Jack's letters is to alter the truth-value of any proposition

¹ *Aristotelian Society Proceedings*, Supplementary Vol. XLV, 1971.

² In 'Personal Identity and Individuation', *Aristotelian Society Proceedings*, 1956-7.

about the past, that is, to make it the case that Jill has written a certain letter when, before Jack wrote, this was not the case. It is not—and this belief is the cornerstone of my argument in *Dr. Who and the Philosophers*—possible to alter the truth-value of any proposition about the past, or of any proposition at all, for that matter. What, we must suppose, Jack's writing his letter accomplishes is to cause it to have been the case that Jill wrote a letter, containing certain expressions, when it was already the case, before Jack wrote, that Jill had written this letter, containing these expressions. This case is like Mr. Dummett's case¹, when a man snaps his fingers before opening an envelope not in order miraculously to change the contents of the envelope from a bill to a letter, nor to change it from being the case that a bill was put in the envelope to its being the case that a letter was put in the envelope, but in order to cause it to have been the case that no bill was put in the envelope when it is already the case that no bill was in the envelope before his fingers were snapped. I cannot make up my mind whether or not this is nonsense, though in my more realistic moments I am inclined to suspect that it is.

What is not nonsense is this. There is no reason why Jack's writing the letter he posts should not be a necessary condition of Jill's receiving either it or one just like it. There is no reason why the date of arrival of Jack's letters should not be predictable (or, perhaps, retrodictable) only from a knowledge of the date on which Jack posts them, and not at all predictable from any knowledge of events which happened prior to their arrival. Jack, though perhaps it is because he is confused, may certainly post his letters in order that Jill may receive them, and so in order that he may receive a reply from Jill. And it may be a true proposition that Jack's posting a letter to Jill on, say, Thursday is a necessary condition of his getting a reply on Saturday. And perhaps the fact that it is already the case that Jill has received the letter which Jack is about to post is no more a reason for denying that Jack is doing something to bring it about that she has received it than is the fact that it is already the case that my aunt will receive the letter I am about to post to her a reason for denying that, in posting it, I am doing something to bring it about that she gets it. Indeed, it is very difficult to find any obvious asymmetry between the two cases, the case of Jack's apparently affecting Jill in the past by what he writes, and my certainly affecting my aunt in the future by what I write, though it is equally difficult to believe that there is no such asymmetry in fact. Perhaps the difference is that the future is unspecified and indeterminate in some way in which the past is not, which might make it possible to bring about what is unspecified,

¹ In 'Can an Effect Precede its Cause', *Aristotelian Society Proceedings*, Supplementary Vol. XXVIII, 1954.

when it is not possible to bring about something which is already completely determinate. But this is a very difficult view to maintain.

The *third* problem, which is raised both by my story about the letter and by *Dr. Who and the Philosophers*, is the problem of the freedom of the will. If we grant, for the sake of argument, that Jill would not have received Jack's letters unless he had posted them, that is, that his posting them is a *necessary* condition of their arrival, it follows, by contraposition, that their arrival is a *sufficient* condition of Jack's posting them. One might be inclined to suppose that, if Jill has already received Jack's letter, and Jill's receiving it is a sufficient condition of Jack's writing it, then it follows that Jack must write it, whether he wants to or not; but this would be a mistake. Though he must write it, in the sense that the fact that Jill has received it and the fact that Jill's receiving it is a sufficient condition of his writing it together entail that he will write it, it may still be the case that he doesn't have to write it; he doesn't have to write it because his wanting to write it is a necessary condition of his writing it, which means that, if he didn't want to write it, he wouldn't. But, of course, if Jack's wanting to write it is a necessary condition of his writing it, it follows that Jack's writing it is a sufficient condition of his wanting to write it, and from this, together with the fact already established, that Jill's receiving a letter from Jack is a sufficient condition of his posting it, it follows that, if Jill has received the letter, Jack must both write it and want to write it. Hence, if we have a case of bringing about the past, we must also have a case of the determination of the act, by which we bring it about, by the past. But is the case so different with any normal act by which we bring about something in the future? Let us suppose that my posting a letter to my aunt is a necessary condition of her receiving one, and that my wanting to post a letter to her is a necessary condition of my posting it. From this it follows, that her receiving it is a sufficient condition both of my posting it, and of my wanting to post it, so, wherever we get a perfectly normal case of bringing about something in the future, we get a case of necessitation of the act whereby I bring that future thing about by the future. I must say that, as a good soft determinist or compatibilist, to use a useful American term, I am not very alarmed by this fact. But then, ought I not to be equally unalarmed by the necessitation of Jack's posting his letter, and his wanting to post it, in 1971 by the fact that it has already arrived in 1851?¹

¹ Based on a story, the author of which I cannot remember.

SYMPTOMS

By JERRY S. CLEGG

I HAVE two aims in this article. First I wish to make a few elementary points about the function of symptoms in a diagnosis. Then I want to attempt to rescue from a clump of confused commentaries an important distinction of Wittgenstein's between symptoms and criteria. I will make no effort to travel farther with that distinction. Brushing the brambles off it is job enough for one analytical expedition.

I

A common way of disparaging patent medicines is to say that they treat only the symptoms of a disease. A sound medicine, so it is often thought, treats the cause of a disorder. This view often has a point, but it invites a cardinal confusion on the relation of symptoms to a bodily disorder. It fosters the notion that a disease just is that which causes symptoms to appear and that the symptoms themselves are mere *signs* of an ailment. But this is generally untrue. A drooping eyelid, to be sure, may be nothing more than a sign of a tumour. Usually, however, symptoms are not just signs of their causes but are, in addition, constitutive of the ailments they are the symptoms of. A headache, fever and sore eyes, for example, are parts of what it is to have a cold, and in their case the ailment just is a set of symptoms. Indeed, if we had no symptoms which bothered us, we would never be ill. We might harbour with no ill effects, for example, a cold virus, as the "Typhoid Mary" harbours a typhoid bacillus; and we would then be the carriers of a disease but not its victims. This truth applies even to the case of a tumour; a drooping eyelid may not trouble us and so not be considered part of an ailment. Still, it is in other symptoms of a tumour that troubles lie. Headaches, blindness and paralysis are the possible effects of a tumour, and it is these symptomatic conditions that we would seek relief from if unfortunate enough to experience them. All by itself, apart from its distressing symptoms, a tumour is no more an ailment than is a bacillus which does not affect its host. Thus, although all symptoms are signs of their causes and although some minor symptoms are no more than that, our ailments are nevertheless constituted by symptoms. It is perfectly proper for this reason to treat symptoms. If we insist that causes should be treated, that is only because we suppose the most efficacious way of getting rid of the symptoms bothering us is to eliminate their source. That supposition may in most cases be true. Certainly a direct treatment of symptoms will often mask rather than eliminate them. Those truths do nothing, however, to undermine the

points that ailments are composed of symptoms and that symptomatic relief is the ultimate aim of all medical treatment.

Because a symptom is generally a part of what we consider a disease, it may function as a means of knowing what is wrong with us. A doctor will observe a sick man's symptoms in diagnosing—often correctly—an ailment. There are obvious limits, however, to the utility of symptoms in establishing a claim to know that a given diagnosis is correct. Some of these limits are trivial; others are of practical significance. The trivial limits consist of the failure of most symptoms to be either sufficient or necessary conditions for having the disease they are symptomatic of. A headache is a symptom of a cold. One need not have a headache, however, to have a cold. Similarly, a headache all by itself doesn't show that we have a cold. Other symptoms have to be present as well. Thus a single symptom will generally fail to be of crucial significance in a diagnosis. This is of little practical importance, for a failure to function as a necessary and sufficient condition does not prevent a symptom from being an important means of identifying what is wrong with a patient. An investigation of symptoms is, indeed, the first stage of almost any competent diagnosis.

A more serious limit on the diagnostic powers of symptoms is that even in association with each other they tend to be ambiguous. A headache, fever and cough, for example, are symptomatic of a host of different ailments—as any reading of a home medical book will show. Thus a diagnosis based on symptoms alone is generally considered to be unconfirmed—that is, a diagnosis which may well be right but which, for all we know, could be wrong.

Because a symptom can be indicative of so many different conditions, we often need something other than symptoms to confirm a diagnosis. This may, but generally will not, be considered a part of a disease. All that is essential is that we find an empirical correlation between a certain set of symptoms which pass for a disease and some phenomena which can serve to rid them of their ambiguity. Chemical tests of a patient's urine, for example, are used to confirm diagnoses of diabetes. What takes place in a test tube is not a part of any ailment, yet a chemical reaction there may serve to establish what ailment a patient is suffering from. Often, in place of tests, it will be the discovered cause of a disease that serves in this capacity. A sick person in whom a typhoid bacillus is found, for example, will generally be held to be suffering from typhoid fever rather than from diphtheria. The bacillus is not itself the disease, for a cause is not what it is the cause of, but it does serve to identify its effects. To distinguish these tests and causal agents from the ambiguous symptoms they clarify, they may—following the usage of Schopenhauer and Wittgenstein—be called criteria.

The function of a criterion in a diagnosis is to eliminate the ambiguity

of the symptoms. Cultures are taken and tests made to check the accuracy of diagnoses based on an observation of symptoms. This can be done without remedying the other limits on the diagnostic powers of symptoms, for a medical criterion need not be either a sufficient or a necessary condition for having a certain disease. The typhoid bacillus whose presence in the key locale, whatever it might be, of a patient's body may be taken as a confirming criterion of typhoid, for example, is *never* a sufficient condition for having typhoid. The victim and the carrier of a disease differ; symptoms always have to be present before we would say anyone was a victim of typhoid. Thus the "Typhoid Mary" does not have typhoid, and the presence of the typhoid bacillus does not prove that she is sick. Even when symptoms are present, indeed, the medical criterion fails to function as a logical proof of a diagnosis. It might be the case, for example, that a carrier of typhoid is sick with diphtheria, but is diagnosed wrongly as having typhoid because the typhoid bacillus is found in his body. Mistakes of this kind are dangerous facts of life. Competent diagnosticians will appreciate them. A diagnosis may be counted as confirmed, but still be wrong. A criterion, then, helps to confirm a diagnosis by eliminating the ambiguity of symptoms, but it does not guarantee anything.

A criterion, too, will often fail as a necessary condition. A doctor may decide, for example, that a patient has diabetes even though he passes standard tests for that disease with no indication that he is a victim of it. Empirical correlations between symptoms and medical criteria will sometimes fail. Thus it is not always essential that a given criterion be found before a diagnosis will be counted as confirmed. Then, too, there are often several criteria for the same disease. There are, for example, four types of animal parasite that cause malaria and that function as criteria for having that ailment. Failure to have a specific type of parasite in one's blood does not eliminate the possibility that one has malaria.

This failure of criteria to overcome all the limits on the power of symptoms as ways of knowing what is wrong with someone does not seriously interfere with their primary function. When an obviously sick person with all the standard symptoms of typhoid is found to harbour the typhoid bacillus, is treated for typhoid, and recovers, we will generally have no reason to suppose that a diagnosis of typhoid was wrong. We can imagine that it was wrong by supposing that all along he had some other disease we failed to identify and that our medication rid him of its cause along with the strain of typhoid bacillus we found, but there is no reason to think these supposings accurate. We will have every right to claim to know that he had typhoid. The criterion can function, then, as a means of knowing and of confirming a diagnosis without becoming a sufficient and necessary condition.

II

In *The Blue Book* (pp. 24–25) Wittgenstein introduces a distinction between symptoms and criteria which follows the practice of diagnosticians. He soon departs from medical examples to apply the distinction to questions of how we come to know about the states of mind of others. His implicit argument is that clarity on how a diagnosis is made and confirmed can contribute to an understanding of how we defend a claim to know about someone else's pains and intentions. The underlying analogy he works with rests on the point that human behaviour—in terms of which we judge states of mind—can seem like a set of medical symptoms in that it may be ambiguous. A groan may, for example, indicate a toothache, disgust, frustration, or a good many other mental states—just as a fever can indicate a variety of diseases. Deciding what a person is thinking or feeling is like deciding what ailment a patient has. Criteria are needed to eliminate the ambiguous possibilities suggested by what may strike us as symptomatic bodily states and behaviour.

Wittgenstein's treatment of the problem of knowledge of other minds in these terms has received intense scrutiny by interpreters of his work. I shall make no effort to resolve all the questions that his analogy between knowledge of diseases and knowledge of mental states raises, for they are manifold. Rather, I shall confine myself to pointing out two sets of rank confusions that have dominated thus far the discussion of his argument and have often led to its unjustified rejection. One set concerns views taken of what he considers a symptom. The other set concerns views taken of what he calls a criterion.

Wittgenstein's example of a symptom is a *throat inflammation* that might be, he remarks, symptomatic of angina. His example of a criterion is a *bacillus* that experience shows us to be associated with the symptoms of angina. Angina itself is 'an inflammation caused by a particular bacillus' (*The Blue Book*, p. 25). In his treatment of the relations of symptoms and criteria to each other and to the ailment of angina he clearly considers the symptom to be constitutive of the anginal disorder. The concept of angina is affected, of course, by the criterion, for not just any inflammation is angina. It has to be 'an inflammation caused by a particular bacillus'. Criteria affect usage and meaning, for with a change in criteria we will change our accounts of what is wrong with a patient. They do not, however, completely determine usage. Whatever the criteria for angina might be or become, angina must remain essentially an inflammation of the throat. Symptoms are always more important than criteria in explaining the use of a disease term. In spite of the diagnostic and textual clarity on these points there has been a high universal confusion among Wittgenstein's interpreters. It has been held that he considers a symptom to be merely empirically associated as a sign with what it is symptomatic of, and that criteria are always, or some-

times, the very same thing as the ailment they are the criteria for. Malcolm, for example, holds that for Wittgenstein 'What makes something into a symptom of y is that experience teaches that it is always or usually associated with y '. For the usual values of y that claim corresponds neither to medical practice nor to Wittgenstein's text. A throat inflammation may be said to be symptomatic of a given bacillus, to be sure, but it is usual to speak of it as a symptom of angina—in which case it is more than empirically associated with what it is symptomatic of. Indeed, when produced by a certain bacillus a symptomatic inflammation just is, as Wittgenstein notes, angina.

Albritton, however, argues in much the same manner as Malcolm, claiming that what is called angina is simply the criterion of angina—while symptoms are empirical signs of the ailment-criterion. Fodor and Chihara follow the line of Albritton and Malcolm on this point, using as an example the roar of a crowd as a possible symptom of a scoring play in a game while a ball going through a hoop could be, at once, both a scoring goal and the criterion of a scoring goal. Wellman has a similar thesis. The criterion of a lemon, he argues, would be features of the lemon that are directly concerned with the meaning of 'lemon'. Symptoms are only indirectly associated by empirical connections established by experience.¹ All these claims invert the relation of symptoms and criteria by erroneously supposing symptoms to be mere signs of what they are the symptoms of. They also amount to misquotations of Wittgenstein's text, for he openly states that medical science might call angina, not a criterion-bacillus, but a symptom-inflammation caused by a particular bacillus. Symptoms, in constituting an ailment, are connected in meaning with the terms denoting the ailment. They are not mere signs of a disorder. Wittgenstein does write that to say that 'A man has angina if the bacillus is found in him' is a 'tautology', but the reference of 'A man' is here to a man sick with the symptoms of angina. That a "Typhoid Mary" does not have typhoid is not a counter-example to his thesis, but an illustration of what his thesis is. Criteria are never just what they are the criteria for. They would be useless as a means of knowing if they were, for, if we do not know what something is, that something is not going to help us come to know what it is.

It may be objected that this account does not take sufficiently to heart Wittgenstein's remark that criteria figure in 'tautologies' and that some criteria are also 'defining' criteria. These are the remarks which have been taken as the prime evidence that Wittgenstein regarded a

¹ References are to the following articles: Norman Malcolm, 'Wittgenstein's Philosophical Investigations', from *Knowledge and Certainty: Essays and Lectures*, Englewood Cliffs, 1963. Rogers Albritton, 'On Wittgenstein's Use of the Term "Criterion"', *The Journal of Philosophy*, Vol. LVI (1959), pp. 845-857. Carl Wellman, 'Wittgenstein's Conception of a Criterion', *The Philosophical Review*, Vol. LXXI (1962), pp. 433-447. C. S. Chihara and J. A. Fodor, 'Operationalism and Ordinary Language: A Critique of Wittgenstein', *American Philosophical Quarterly*, Vol. II (1965), pp. 281-295.

criterion as equivalent to what it is the criterion for. Aside from being patently disastrous in its implications and in conflict with Wittgenstein's point that angina is an inflammation, not a criterion-bacillus, this judgement is unwarranted for still other reasons.

Wittgenstein's point is that the question 'How do you know this man has angina?' is not stopped by an appeal to the presence of symptoms alone. It is not stopped entirely by an appeal to the criterion-bacillus either, for it is imaginable that someone should be the carrier of a disease while being the victim of another. As a practical matter, however, the question is normally stopped, for we have come down, Wittgenstein notes, not to a question of fact but to a 'convention' as to what a disorder is to be called. In acknowledgement of the conventional role of criteria, Wittgenstein characterizes them as functioning in 'definitions' and 'tautologies'. His characterization, however, is to be read in light of his views on the epistemological status of definitions. Immediately after his discussions of symptoms and criteria in *The Blue Book* he distinguishes between 'explanations' of usage which give a full account of a word's meaning and ornamental 'definitions' which are of little philosophical value—however useful they may be in special contexts.

This distinction between 'explanation' and 'definition'—together with the injunction to pursue the first and to avoid a fixed concern with the second—may require a lot of explication, but it is clear that it would be a totally pointless distinction if Wittgenstein regarded a 'definition', 'convention' or 'tautology' as the full conceptual equivalent of a term. There is, then, no need to prune Wittgenstein's text to rid it of an embarrassing, unworkable equation of criteria with what they are the criteria for. Those who read Wittgenstein in a way that makes this necessary are the ones who have failed to take due note of his quarrel with the practice of trying to resolve a philosophical question with the help of a mere definition.

The confusions over the relation of symptoms to a disorder is extended by all these writers to a series of differing confusions on how criteria function in confirming a judgment. In his critique of Wittgenstein Albritton supposes that in *The Blue Book* a criterion is a necessary and sufficient condition for an expression applying to some object—as a bird cage just is, on occasion, sufficient to be one's baggage. But criteria do not function in this capacity. By themselves they settle nothing, and even in association with symptoms they do not guarantee the truth of a diagnosis. Wittgenstein never says that a bacillus is a disease, and in a relevant passage in *The Philosophical Investigations* (I 84) he points out that we can imagine someone *always* doubting that even the simplest beliefs are right. Criteria affect a concept but without determining in a demonstrable way that it applies.

Malcolm holds that criteria need not be sufficient conditions but that they do guarantee that an expression will apply sometimes. If someone always had endless doubts about the application of some term, he would have no criterion, Malcolm claims, for the use of that term. But this distorts the features of criteria, is in no sense true, and is not a derivative of what Wittgenstein says. Medical criteria serve to confirm a particular diagnosis. They are not appealed to in order to show that someone at sometime must have been right in saying, for example, that a patient had typhoid. They are looked for to confirm a specific diagnosis. They would be useless otherwise, for we want to know with their help if we have a certain disease—not if the possibilities that suggest themselves have been realized in times past and places distant. The confirmation they offer, however, is not beyond all challenge. We can imagine, as Wittgenstein suggests, someone always doubting the truth of a diagnosis. We can, too, easily imagine that as medical science advances old diagnoses will be abandoned. It may turn out, for example, that the present distinction between serum and infectious hepatitis will not be used in a few years. It might be decided that serum hepatitis is not a distinct disease. If that were decided, a medical technician of the future would in no sense be obliged to admit that at least some present diagnoses of serum hepatitis—using present criteria—were sound. A criterion does not guarantee anything—not even that a disease term must apply at least sometimes. They are both stronger and weaker than that. There is no cause to suppose that Wittgenstein ignores this facet of the medical use of ‘criterion’. Malcolm’s claim that a word cannot always be used wrongly is a dubious invention of his own.

Wellman’s critical contentions on the function of criteria are confused for allied reasons. He holds that the criterion of a word ties it to its reference, as ‘lemon’ is tied to a lemon by its meaning. But this is utterly absurd. The question ‘Is this man suffering from angina?’ is very different from the question ‘Is this the anginal bacillus?’. The first question is to be answered by undertaking a diagnosis complete, perhaps, with tests. No diagnosis is relevant to the second question. If it were, a regress would be involved in all diagnoses, for we would have to diagnose what we use to justify a diagnosis. We acquire the ability to use descriptive expressions through training, Wittgenstein argues. How do I know something is red? Because I’ve learned English (*The Philosophical Investigations* I 381). How do I know something is a lemon? For the same reason—not because I have criteria for its diagnosis. How do I know that something is an anginal bacillus? If I were to know, it would be because I had been trained as a medical technician. Criteria which are useful in confirming a diagnosis are not the means by which words are generally connected to their references. There are no symptoms of a lemon, of a bird cage, or of a scoring play in a game. There are

no criteria for these things either, for criteria function only in conjunction with symptoms.

Fodor and Chihara admit this point to a certain degree, claiming that for Wittgenstein only the usage of some words have to be taught with the help of criteria. But this is false *even of medical terms* where criteria are clearly present. It is also untrue to Wittgenstein's text. In undertaking a diagnosis one must have already learned a language. The problem a criterion solves is one that arises from ambiguity. We can reasonably suppose in many cases of illness that different diseases are involved. The question to be answered is which of the two or more plausible diagnoses is right. It is not the problem of teaching or learning the use of a word. Word usage is already mastered and the problem remains, for we don't know if a word we fully understand is the one to apply in this case. A doctor is apt to know full well what 'malaria' means and to be capable of explaining its use. What he may not know is if a given patient has malaria. Diagnostic skills are not acquired with the mere learning of a language. If they were the time required for gaining a medical degree could be cut drastically. Besides, the criteria used in diagnoses are not, as a matter of fact, used in most teaching situations. I know what angina is, but I've never seen an anginal bacillus and would not recognize one if I did. Mastery of the technician's microscope is not required for most teaching purposes. As Wittgenstein strenuously insists in connection with the function of criteria, we do not learn a language by means of strict rules comparable to those of the diagnostician (*The Blue Book*, p. 25). Criteria are not primarily means of teaching. They may instruct, but their chief function is to confirm a surmise.

III

The summary of the argument may now stand as the conclusion of this paper. Wittgenstein's analysis of the problem of knowledge of other minds rests on an analogy between how we decide what ailment a sick person has and how we decide what others think and feel. His crucial distinction between symptoms and criteria is a distinction to be understood in diagnostic terms. Lemons, bird cages, and basketball hoops illustrate nothing about his thesis. These are matters that have not been appreciated by a significant range of his interpreters who have carried common confusions on the rôle of symptoms into their critiques and expositions of Wittgenstein. If attention is paid to how symptoms do operate in a diagnosis, the rôle of antithetical criteria as eliminators of ambiguity can be seen. When that is seen, it becomes clear that criteria are useful means of confirming surmises but that they guarantee the truth of nothing. They do not prove a given diagnosis beyond all possibility of doubt, nor the proposition that at least some diagnosis, made at some time, is accurate. They need not be a part of the phenomenon

they are the criteria for, and they are never just the whole of that phenomenon. They also have nothing to do with tying common descriptive predicates to their objects, and they are not essential to the teaching of a language. They serve to adjust or make more precise concepts we have already mastered, and they do that by providing a way of confirming diagnoses made by linguistically competent persons on the basis of ambiguous symptoms that are constitutive of what they are the symptoms of. These are all points fully compatible with what Wittgenstein has to say. If they are kept in mind his analysis of the problem of knowledge of other minds may be approached as the interesting and promising analysis it is. Failure to keep them in mind can only leave his work lost in the thicket of muddles and distorted readings that have grown up around it.

Mills College, Oakland, California

INDIFFERENCE AND VOLUNTARINESS

By JAMES B. BRADY

IN her article "The Range of Application of "Voluntary", "Not Voluntary" and "Involuntary" (ANALYSIS, Vol. 26, April, 1966, pp. 149-152), Lorene Gordon argues that ordinary actions are voluntary only where ordinary action is taken to mean intentional action. If the action is not 'intentional', meaning, according to Gordon, 'what the agent is doing from his own point of view', 'voluntary' and 'not voluntary' do not apply. I am interested primarily in her further claim that in certain cases where actions may be called intentional, 'voluntary' and 'not voluntary' do not apply. The example which she adduces for this position is the following:

... I am holding a book in my hand when I see that a child in the room is about to burn himself on the open fire. I drop the book and move to prevent the injury. Is my dropping of the book intentional? voluntary? ... I don't *mind* that I'm going to drop it and so the question as to whether or not I could, or even would, do otherwise simply does not arise. Despite my knowing what I'm doing and not minding, it seems perfectly reasonable here to suggest that the concepts 'voluntary' or 'not voluntary' simply do not apply (p. 151).

Gordon's argument is that it is the fact that the agent doesn't care or doesn't mind dropping the book which makes 'voluntary' and 'not voluntary' inapplicable. We only have to change the example, however, to a situation where the action in question has some significance to see that the concepts do apply. Suppose that a workman is carrying a heavy timber on the roof of a building. Upon hearing the call to lunch he

drops the timber. In this example the workman knows that he is dropping the timber and does not mind dropping it. With the element of risk of injury to a pedestrian on the walk below, would there be any hesitation in saying that 'voluntary' and 'not voluntary' apply to this action regardless of the fact that he just didn't care whether he dropped it?

As I have described it the case of the workman is an example of acting recklessly; by acting recklessly I mean acting with the awareness of a substantial and unjustifiable risk. (The actual awareness of the risk distinguishes recklessness from negligence; in acting negligently the agent is inadvertent to the risk. The element of indifference distinguishes recklessness from acting purposely.) Can the workman in the example deny his culpability by maintaining: 'I didn't mind dropping the timber. So the question as to whether I could, or even would, do otherwise simply doesn't arise . . . I just didn't care'? In both intentional and reckless action there is awareness of what one is doing. The capacity to control conduct, therefore, would seem to be the same in each. The mere fact that the agent didn't care about what he was doing doesn't rule out the question of whether he could have done otherwise. Recklessness (which implies indifference) is an instance, therefore, where 'not caring' does not imply that the action is neither 'voluntary' nor 'not voluntary'.

I agree with Gordon that it would be odd to ask whether the person in her example could have done otherwise. The reason it would be odd, however, is not that the person didn't care about dropping the book but the fact that we have no reason to ask whether he could have done otherwise since the action is socially insignificant. In my counter-example, where the agent is also indifferent to what he is doing, the fact that the action involved risk of harm raises the question of whether he could have done otherwise. The mistake which Gordon makes is rather similar to the one pointed out by Schiffer (*ANALYSIS*, Vol. 25 (Supp., No. 3), January, 1965, pp. 94-98) in his criticism of Cavell's claim that normal everyday actions are neither voluntary nor not voluntary. According to Schiffer it would be puzzling if one were asked whether a normal everyday action (one in which there is nothing "fishy" or peculiar) were voluntary. Why it is puzzling, however, is that since normal everyday actions *are* voluntary there is no reason for asking the question. In Gordon's example there is no reason to ask whether the action is voluntary since it is hard to imagine why anyone would ever care whether or not it is voluntary. Nothing is at stake. In my counter-example we really are interested in whether the action was voluntary since the question of whether he is to blame for causing risk hinges on this.

MANNISON ON DOING SOMETHING ON PURPOSE BUT NOT INTENTIONALLY

By G. A. SPANGLER

TAKING up a suggestion made by Austin in 'Three Ways of Spilling Ink', D. S. Mannison has tried (ANALYSIS, December 1969) to construct a case in which a person does something on purpose but not intentionally. I want to argue here for the modest claim that Mannison has not found such a case.

Consider the situation Mannison presents. A man is driving his car down a twisting mountain road when his brakes fail, letting the car roll uncontrollably fast toward a sharp curve and certain death for its passenger. On his left is a tavern owned by an aggressive lout who has frequently assaulted his brother-in-law, the owner of a gas station on his right. Calculating that he can save himself by smashing into one of these establishments, he turns his car into the tavern.

Mannison first argues that smashing into the tavern was something done on purpose. Because he could have chosen to drive his car into the gas station instead of smashing into the tavern, the situation presented 'non-trivial alternatives' to the driver, i.e., the only alternatives were *not* 'to do X' or 'not to do X'. And since 'failure to smash into the tavern would constitute not doing what has been chosen', there is 'no reason to deny that smashing into the tavern was done *on purpose*'. What is crucial in this case is the element of choice. After having crashed his car into the tavern the driver could not claim that he 'had no choice' nor that he had done what he did 'simply for the purpose of saving his life'.

But suppose there were a passenger in the car who, after the driver had chosen to smash his car into the tavern, asked whether the driver intended to hit the tavern. Having deliberately chosen the tavern as his target, the driver could not honestly answer with a 'no'. He may not have intended to hit the tavern when he started down the hill, but he intended to do so after he had chosen to do so. Mannison says an objection like this one need not be considered because he is not employing a version of an argument Austin had used concerning a case in which a man is ruined 'as a consequence' of being compelled to pay a debt. But Mannison shows that Austin's case differs from his own in at least one important respect: smashing into the tavern was not a *consequence* of doing something else. So Austin's argument is irrelevant to the case at issue, and the objection I raise stands even if Mannison cannot take and does not choose to try Austin's way out.

But there is a more serious irrelevancy in Mannison's argument. It consists of showing that smashing into the tavern is neither necessary nor sufficient for saving the driver's life. Then, since success at smashing

into the tavern is compatible with failing to save one's life by smashing into the tavern, it is supposed to be apparent that smashing into the tavern was done on purpose but not intentionally. But this is like arguing that a person who assaults another on purpose and with intent to do murder does not assault his victim intentionally if he fails to murder him. Assault is not a necessary condition for murder, but whether it is or not is irrelevant in cases like this: if a thug fails to assault his victim, he *eo ipso* does not assault him on purpose. And it is likewise an irrelevant consideration that assault is not a sufficient condition for murder: if a thug assaults his victim with the intention of murdering him, but fails to murder him, he does not succeed in doing one of the things he intended to do. But he wouldn't get very far in court by arguing that, for this reason, he hadn't intended to assault his victim. Just so, it would be right (fair) to accuse the driver of smashing his car into the tavern intentionally, even though he was trying to save his life.

University of Alberta

TRUE PROPOSITIONS: A REPLY TO C. J. F. WILLIAMS

By CHARLES SAYWARD

WHEN Williams originally proposed (9) as a definition of '— is true' it looked as if '— is true' could be read as '— is a true sentence' or '— is a true proposition', depending on whether the name of a sentence or proposition replaced the blank. For names of sentences apparently could replace the blank in (9) just as could designations of propositions. But by proposing (24) or (25) in place of (9) Williams must be using '— is true' to abbreviate '— is a true proposition' (a practice I shall adopt in this paper). Names of sentences cannot replace the blanks in (24) or (25). (See ANALYSIS 29.4 and 32.2.)

In his original paper (p. 122) Williams proposed (9) as an *explicit* definition of '— is true'; so apparently the same is being proposed of (24) or (25). As I understand it such a definition should allow one to replace the defined expression in any sentence in which it is used in the relevant sense by the defining expression without change of truth-value. So if we are to accept Williams' proposal, we should be able to put (24) or (25) for '— is true' in any sentence in which it is used without change of truth-value. This is a very strong condition which I do not think can be met.

Consider the true sentence

- (34) Every proposition expressed by a true eternal sentence of English is true.

Replacing '— is true' in (34) by (24) and (25) yields these two sentences:

- (35) $(\exists p)$ [(If anyone were to assert every proposition expressed by a true eternal sentence of English, he would thereby be asserting that p) & p]
 (36) $(\exists p)$ [(Every proposition expressed by a true eternal sentence of English is the same proposition as the proposition that p) & p].

I see no way of testing the truth of such sentences as (35) and (36) without appealing to the following characterization: (A) A sentence of the form ' $(\exists p) \dots p \dots$ ' is a true sentence just in case some substitution instance of ' $\dots p \dots$ ' is a true sentence. Using (A) it is clear that both (35) and (36) are false. No replacement of ' p ' will convert the open sentences, obtained by removing the quantifier, into true sentences.

It is worth noting that sentence (34) says the same as

- (37) (x) (x is a proposition expressed by a true eternal sentence of English $\supset x$ is true).

Replacing '— is true' by (24) or (25) in (37) does not result in a change of truth-value. But this fact does nothing to show that either (24) or (25) meet the condition for being adequate as explicit definitions. It may be that either of them can replace '— is true' in just those sentences in which the blank is replaced by a singular term or name variable without change in truth-value. If this more moderate condition holds, either (24) or (25) could be used to provide an adequate contextual definition of '— is true'. In this case Williams' analysis would still have much explanatory value.

Of course, for this condition to hold and, indeed, for one to even recognize (24) and (25) as meaningful predicates the practice of binding sentence letters must be legitimate. I think Williams needs to say more by way of justifying this practice.

I am willing to grant that the argument form 'A says that p , $p=q$, \therefore A says that q ' is valid. It follows that as long as it is legitimate to bind sentence letters at all, there are not special facts about 'A says that' which vitiate Williams' analysis. But we need an argument to show that it is legitimate to bind sentence letters at all.

There are a number of sentences of the form ' $(\exists p) \dots p \dots$ ' and ' $(p) \dots p \dots$ ' that cannot be translated into English without using '— is true' or a variant. Williams has not provided such a translation of

S(8'), for example; and I doubt that one can be provided. It does not follow that S(8') is unintelligible. As long as (A) is an acceptable characterization of the truth conditions of such sentences there is no reason for questioning their intelligibility. This is just something I overlooked in my original reply to Williams.

The characterization (A) is an immediate consequence of the following characterization of quantification in general: (B) An existentially quantified sentence is true just in case some substitution instance of the open sentence obtained by deleting the initial quantifier is a true sentence. Williams cannot use (B) to justify (A) because it is controversial whether (B) is acceptable. Moreover, considerations analogous to those which may prove to vitiate (B) may prove to vitiate (A) as well. For example, John Wallace ('Convention T and Substitutional Quantification', *Nous*, 5.2, pp. 199-211) has recently argued that no semantical theory of elementary number theory can employ (B) without violating Tarski's Convention T unless the theory is so strengthened as to make it a mere notational variant of one which employs the referential account of quantification. Suppose this conclusion is correct. A similar conclusion might well hold for the semantics of any theory in which the treatment of sentence letters as bindable variables is essential. (In *Philosophy of Logic*, [Englewood Cliffs, New Jersey: Prentice Hall, 1970, pp. 74-75] Quine describes languages in which this practice might be a mere convenience.) For such a theory it may well be the case that if the semantics of the theory utilizes (A) it will violate Convention T. Unlike the case with name variables, there is no obvious alternative to (A). So it may be that the semantics of the theory must violate Convention T.

I do not know if this conclusion can be sustained. Until the semantics and syntax of the sort of theory envisaged is spelled out it is anybody's guess. I have just tried to make out the case that it is not all that obvious that (A) is acceptable. Its acceptability has to be established before one can use either (24) or (25) as an analysis of '— is true'. For without appealing to (A) I see no way of making sense of sentences resulting from applying (24) or (25) to singular terms that refer to propositions.

University of Nebraska

TRUTH AND PREDICATION

By O. R. JONES

IN his composite rejoinder on the topic of truth (ANALYSIS, Dec. 1971), C. J. F. Williams makes it clear that he considers what

(1) $(\exists p) [(A\text{'s statement states that } p) \ \& \ p]$

says about A's statement to be analogous to what

(2) $(\exists x) [(x \text{ is a man} \ \& \ \text{Edith married } x) \ \& \ \text{Alice married } x]$

says about Edith. Let us see where these two are analogous before going on to consider the extent to which they are disanalogous. In order to show that they are analogous we have to invoke a principle of predication according to which a proposition is, as a whole, about an item if the introduction of a subject-expression referring to that item completes what would fail to be a proposition without that, or some other comparable, subject-expression. For convenience, I will henceforth refer to this principle as 'the contextual principle'. According to the contextual principle, all we have to do in order to expose what (1), as a whole, says about A's statement is to take out the expression 'A's statement', and similarly we may expose what (2) says about Edith by taking her name out. The principle lets us believe that (1), as a whole, is as much about A's statement as (2), as a whole, is about Edith.

It is by the same principle that Williams decides that

(3) A's statement states that the door is closed, and the door is closed

is, as a whole, about A's statement. So it is a principle which allows predicates to run indiscriminately over logical connectives. It is worth highlighting this peculiarity of the principle by attending to some paradoxical results that can be obtained by applying it. It will require us to say, for example, that

(4) John is tall, and Mary is married

says something false about John in the case where John is tall, but Mary is single. As against this it would seem natural to say that all that (4) says about John is that he is tall, which happens to be true, and that nothing more is said about John by adding that Mary is married. There would appear to be a sense of the word 'about' in which what I have said in that last sentence is true. It is the sense that we have in mind when we ask: 'Tell me more about John', and refuse to take 'Mary is married' for an answer. Another paradoxical result of applying the principle is that one has to say that, if John is tall,

(5) John is tall or Mary is married

says something true about Mary even though she is single. Surely there is a sense of the word 'about' in which (5) does not say anything true about Mary.

The difference between (4) and (5), on the one hand, and examples like

(6) John is tall,

(7) John is taller than Mary,

and

(8) John is tall and he is bright,

on the other, is that (4) and (5) are truth-functions of a proposition which is not about John in any sense. John would have to be no different for the second conjunct in (4), and the second disjunct in (5), to be true from what he would have to be if these were false. Clearly, (6), (7) and (8) are not dependent on the truth or falsity of such a proposition.

Now consider an example which superficially appears to be like (4):

(9) John is the father of George and George is the father of Fred.

This example is comparable to (4) in that it is conjunctive and in that the second conjunct is a proposition that would not, on its own, say anything about John. However, it is obvious that since both conjuncts are also about George the predicate of the second conjunct links up with the predicate of the first conjunct in such a way that more is said about John than would be said about him by the first conjunct alone. The augmentation which accrues to the predicate of the first conjunct in virtue of its link-up with the second conjunct is captured in

(9a) John is the father of the father of Fred,

which is equivalent to

(9b) John is the paternal grandfather of Fred

and is entailed by (9). Though the conjuncts in this example happen to be relational ones which have the common subject 'George', it need not happen that way in order for there to be a linkage of predicates, as is obvious in the example

(10) John is a grocer and a grocer is a shopkeeper

which entails

(10a) John is a shopkeeper.

In an attempt to sort out some relevant differences here let us call that which results from the application of the contextual principle a 'context' (cf. A. N. Prior in *Objects of Thought*, p. 150). I think we have four different kinds of contexts represented in our examples. First, there is the type represented in (6) and (7) which do not include a sentential connective; I shall call it a 'predicate context'. Secondly, there is the type represented in (8) where we have a connective linking two conjuncts in such a way that they have the same subjects; let us call this the 'linked-subject context'. Thirdly, we have cases like (9) and (10) where the connective effects a link between the predicates, as already described; so this may be called the 'linked-predicate context'. Finally, we have cases like (4) and (5) where there is neither linkage of subjects nor linkage of predicates, but simply a linkage of propositions; and I shall call this the 'linked-proposition context'. The linked-proposition context is capable of analysis in the obvious way. The linked-proposition context of (4) for example, consists of a predicate context linked by the conjunction connective with a complete proposition which, on its own, would say nothing about John in any sense of 'about'.

I think it is fairly clear that (3) is a linked-proposition context. The subject of the first conjunct is clearly different from the subject of the second conjunct, and I cannot imagine anyone wanting to argue that there is a linkage of predicates via an entailment comparable to what we have in the case of (9) and (10). Williams, however holds that (3) entails (1), and also holds that

(2a) John is a man and Edith married John, and Alice married John entails (2), which he regards as analogous to (1). Let us look at (2) for moment. The part preceding the second ampersand in (2) entails

(2b) Edith married a man,

so there is a linkage of predicates there. It is easy to see that the remaining predicate is also linked-up, and again we have a shorthand way of exhibiting the augmented predicable which is applied in this case to Edith. It is

(2c) ———married Alice's husband.

Can (1) do for (3) what (2) does for (2a)? Can it show that we have a linked-predicate context in (3)? I do not think so; the least that we would need to serve that purpose is

(1a) $(\exists p) [(A's \text{ statement states that } p) \ \& \ p \text{ is true}]$,

and what one might hope to show with the help of this is that

(3a) Λ 's statement states that the door is closed, and it is true that the door is closed

is a linked-predicate context. However, (1) has to make good sense quite independently of (1a) since (1) is supposed to contain an explicit definition of '— is true'. So reference to (2) is only to the point if there is an analogy between it and (1) as it stands, and I fail to see what analogy there is beyond the fact that in both cases there are contexts. The context of 'A's statement' in (1) is, if anything, a linked-proposition context; at any rate it is something like (3) that will make it true, and (3) has not been shown to be anything other than a linked-proposition context.

At this stage, it will help to bring Williams's position into perspective if we contrast it with Austin's. Austin uses the word 'about' without explicitly distinguishing different senses of the word. He suggests that what words are used to communicate about may be called 'the world'.¹ Thus, he would regard (6), for example, as saying something about "the world" (or, as he sometimes seems to think, a part of it—the situation which consists in John's being tall perhaps). He would also regard

(6a) It is true that John is tall

as being about "the world", but, furthermore, he regards a proposition like (6a) as being about what he would call a statement as well, about the statement that John is tall, and that statement itself he regards as a part of "the world". (Austin writes: "TstST refers to the world or any part of it inclusive of tstS.")¹ Thus, he would suppose (6a) to be about both the statement that John is tall and about John's being tall (which, it seems, he was prepared to call 'the world', or sometimes 'part of the world', or 'a fact'). Finally, Austin thought that what a proposition (or statement as he would say) like (6a) says about a statement—(6) in this case—and about the world, fact or whatever, is that they are related in a certain way. His version of the correspondence theory is an attempt to explain what the relation is. Clearly, Austin regards (6a) as a relational statement like, for example,

(11) This score corresponds to the music you hear now,

and, for our purposes, this falls into the same category as (7). The context here is a predicate context with no connective involved. Austin's position is in stark contrast with Williams's, though both think that

(12) A's statement is true

says something about A's statement. According to Austin, the analysis of the context found in (12) is itself a predicate context, and there is no question of going on to show that it is something else. By contrast, Williams's analysis depicts it as being something different from a plain

¹ *Truth*, (ed.) G. Pitcher (Prentice-Hall: 1964), p. 26.

predicate context. Since both put their view in terms of (12)'s being a statement about a statement, it is worth drawing attention to the extent of the difference.

However, Williams does not particularly wish to be associated with Austin's view, but with Ramsey's, so I now want to evaluate Williams's analysis by setting it in contrast with Ramsey's. Consider what Ramsey says about

(13) For all p , if he asserts p , p is true.

Ramsey writes: 'We have in English to add "is true" to give the sentence a verb, forgetting that " p " already contains a (variable) verb' (Pitcher, p. 17). He is arguing that what holds of 'It is true that Caesar is murdered' also holds of (13), namely, that 'It is true that' in the one case, and 'is true' in the other, are logically superfluous. He regards 'is true' as explicable in the light of 'It is true that'; the logical status of the verb-phrase is to be understood in the light of the adverbial-phrase, and not the other way round. Another way¹ of putting it would be to say that 'It is true that' is an operator for forming a sentence out of a sentence. In particular, it is the identity truth-function operator, and that is why Ramsey thinks that, so far as the logic of the matter is concerned, it can be dispensed with. This is the very heart of what he wants to say about truth. This is also why he thinks that

(14) For all a, R, b , if he asserts aRb , then aRb

is better than (13); with respect to (14) he says that 'is true' would be 'an obviously superfluous addition'. This, he thinks, is not quite so obvious in the case of (13) because English grammar requires us to add 'is true' to give the sentence a verb. That is to say, 'is true' in (13) is only a verb by the requirements of English grammar. If we want to grasp its logical significance then we should look back to the adverbial 'It is true that'.

Ramsey's move from (13) to (14) suggests that he would not be happy with Williams's method of quantifying over propositions. He would, perhaps prefer

(15) For some a, R, b , both A's statement states that aRb and aRb

to Williams's (1), but if he were happy with (15) he really should be equally happy with (1). If he were, then I think he would welcome it on grounds altogether different from Williams's. He would welcome it, not as giving a context which was an explicit definition of '— is true', but as providing a context which shows how '— is true' can harmlessly

¹ I am very grateful to my colleague, Mr. Peter Smith, for drawing my attention to this, and also for many discussions of earlier drafts of this reply. He is in no way responsible for any faults that may remain.

drop out. (1) should be understood as providing for 'A's statement' a linked-proposition context such that (1) is true if there is a proposition like (3) which is true of A's statement, a proposition where the context clearly consists of a predicate context conjunctively linked with a proposition. That is the proposition which is said to be true, and it is also the proposition from which the grammatical predicate '— is true' has been harmlessly detached. The important business of elucidating truth, as Ramsey sees it, comes after all this has been done, and it is necessary to show why this can be done. The answer lies in grasping the significance of 'It is true that', which is not itself one of the four contexts that I have mentioned, but a particular kind of operator. Far from conceding to Williams that the so-called first theory (the redundancy theory usually attributed to Ramsey) is 'less subtle' than the second 'more powerful' theory, Ramsey would hold that the second theory is simply an aid to show how some troublesome cases turn out to be amenable to the treatment of the first theory.

The conflict between Williams's view and Ramsey's can be brought out very effectively by considering what each would say of a case like

(16) It is true that the house an old one.

Williams takes the view that the predicable '— is true' occurs in 'It is true that', and that we have an explicit definition of that predicable in

(17) $(\exists p) [(— \text{ states that } p) \ \& \ p]$.

So we can move from (16) to

(16a) That the house is an old one is true,

to get a designating phrase, and thence to

(16b) $(\exists p) [(That \text{ the house is an old one states that } p) \ \& \ p]$.

This is supposed to be primary and elucidatory; we are supposed to understand better what is said when something is said to be true if we realize that the predicable can be put in the form of (17). However, I do not think that Ramsey would have it that way at all.

Our (16b) is true if

(16c) That the house is an old one states that the house is an old one, and the house is an old one

is true, and that is true if

(16d) The house is an old one

is true. The important difference between (16d) and (16) is that 'It is true that' has fallen off, and, thanks to Ramsey's redundancy theory, we know why. But of course the bloating up of (16) into (16b) was

quite unnecessary and unilluminating, for what we need is to realize what kind of operator 'It is true that' is, and thereby why (16) is equivalent to (16d). With an example like (12) there is indeed some point in a move such as the one Williams advocates, but it does not constitute a second theory. It is merely a move that helps to show how the original theory applies to rather unlikely looking cases. Let us give Ramsey the last word: 'When all forms of proposition are included the analysis is more complicated but not essentially different'.

University College of Wales, Aberystwyth

BLACK AND THE INDUCTIVE JUSTIFICATION OF INDUCTION

By BREDO C. JOHNSEN

RECENTLY, Professor Max Black has maintained once again that the argument that an inductive justification of induction is possible without circularity 'has not yet . . . been shown to be mistaken'.¹ He does not, however, make the stronger claim that the argument has been *shown not* to be mistaken, and there is some reason to think that this is not the result of an oversight on his part; after wide experience, he is willing to concede that competent philosophers may honestly disagree on this issue. In fact, his pessimism about the possibility of settling it decisively led him some time ago to rest it on another, presumably clearer, issue—namely, the possibility of increasing the strength of first-level inductive arguments by means of second-level arguments which make use of the very same rule as do the arguments they strengthen.²

I will show here both that Black's argument for the existence of this possibility fails decisively and that in fact the possibility does not exist. Thus if he is justified in allowing the issue of circularity to rest on this other issue, the issue of circularity is decided against him.

A preliminary task is to set out the particular second-order inductive argument and the particular inductive rule in terms of which Black states his own argument, as well as his characterizations of two crucial terms.

¹ See 'Induction' in *Encyclopedia of Philosophy*, ed. Paul Edwards.

² See 'Self-Supporting Inductive Arguments', in Max Black, *Models and Metaphors* (Ithaca, N.Y., 1962), pp. 209–218. All subsequent page references are to this paper.

The second-order argument is the following:

(a): In most instances of the use of R in arguments with true premises examined in a wide variety of conditions, R has been successful. Hence (probably): In the next instance to be encountered of the use of R in an argument with a true premise, R will be successful (218).

The rule is the following:

R: To argue from 'Most instances of *A*'s examined in a wide variety of conditions have been *B*' to (probably) 'The next *A* to be encountered will be *B*' (218).

The two crucial terms are characterized as follows:

Inductive arguments governed by R vary in "strength" according to the number and variety of favourable instances reported in the premise (211).

... at any time in the history of the employment of an inductive rule it has what may be called a degree of reliability depending upon its ratio of success in previous applications (212).

Black states his thesis succinctly:

Now in claiming that the second-order argument (a) supports the rule R, I am claiming that the argument raises the degree of reliability of the rule, and hence the strength of the arguments in which it will be used; ... (212).

It is a simple matter to show that this thesis cannot possibly be true. The degree of reliability of a rule at a given time in the history of its employment is a function of its ratio of success in previous applications (see characterization above). Thus at a given time the value of this function for R is, say, m/n . Now Black's claim is that we can use (a) to *raise* the degree of reliability of R to some higher figure, m'/n . But the premise of (a) *simply states the past ratio of success of R*, which was the basis of the original determination of R's degree of reliability. Apparently we are to understand that (a) provides a second way of determining that degree of reliability, on the basis of the very same data, and with different results.

This is absurd. Either the degree of reliability of a rule at a given time is a function of its ratio of success up to that point, or it is not. If it is, then for some m and some n , m/n is *the* measure of that degree of reliability, and it simply makes no sense to talk of raising or otherwise altering it. (The fact that a rule's degree of reliability is subject to fluctuation over time is of course not to the point here.)

There is a second, equally extreme difficulty with the argument. Even if Black were completely clear about the notion of reliability, he could not consistently maintain that the reliability of an argument's rule bears on the argument's strength. (Recall that his thesis is that one '*... raises the degree of reliability of the rule, and hence the strength of the argument. . . .*' My italics.) The strength of an argument varies

according to the number and variety of favourable instances reported in the premises, and pretty clearly this is supposed to mean that it varies solely with those two factors (see characterization above). If so, then the possibility of increasing the strength of first-order arguments by use of Black's second-order arguments is ruled out; for the latter are concerned exclusively with the reliability of the rule in question, and the strength of arguments is not, even partly, a function of the reliability of their rules.

It might occur to one here that Black has simply spoken carelessly in characterizing the strength of arguments, since it might seem terribly implausible that the strength of an argument is completely independent of the degree of reliability of its rule. But in fact their independence is absolutely essential to Black's argument. Otherwise, he would be committed to the view that one must take the rule's degree of reliability into account in determining the strength of an argument, but that having done so, one can *redetermine* it (with a different result) by appealing a *second* time to the rule's degree of reliability (and nothing else) through the use of (a). This remains absurd.

Thus, on his own understanding of the notions of reliability and strength, both of Black's crucial claims are false: (i) that argument (a) can be used to alter the reliability of a rule, and (ii) that the strength of an argument can be increased by an increase in the degree of reliability of its rule. I suspect that Black is at least roughly right in his understanding of the two notions, but whether or not that suspicion is correct, it is certain that *if* he is right about them, his argument fails.

Finally, then, if the possibility of an inductive justification of induction depends on the possibility of increasing the strength of inductive arguments by means of second-level arguments making use of the same inductive rule, then the former possibility does not exist.

University of Houston

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, R. G. Swinburne; J. H. Benson, Miss M. A. Boden, Alec Fisher, Andrew Harrison, R. F. Holland, Stephan Körner, A. C. MacIntyre, A. M. MacIver, Bernard Mayo, A. R. White, Peter Winch. This committee is responsible for appointing and advising the editor and for the general policy of the paper.

SUBSCRIPTIONS. The subscription to ANALYSIS, Vol. 32, is £1.50 for individuals in Great Britain, £1.75 for individuals abroad, and £2.00 for all institutions, post free. Each volume comprises six numbers, five of 32 pages and one of 48 pages, appearing within the academic year—in October, December, January, March, April and June. Orders should be sent to Basil Blackwell, 49 Broad Street, Oxford, or placed with any bookseller.

CONTRIBUTIONS. Articles submitted for publication should be addressed to Dr. C. J. F. Williams, 1 Fossefield Road, Midsomer Norton, Bath, BA3 4AS. Contributors are asked to note the following requirements.

Articles should normally not exceed 3,000 words in length. Occasionally, however, longer contributions can be accepted;

They must be typewritten in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgement is required a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope and international reply coupons of the requisite value, whether for air or surface mail.

Proofs are normally read by the Editor. A contributor who wishes to read his own proofs should notify the Editor at the same time as the MS. is submitted.

OFFPRINTS. Contributors will be supplied free of charge with a limited number of offprints of their articles. Further copies may be ordered at a charge from the publishers. Orders for these, if required, should be placed direct with the publishers when an article has been accepted for publication.

The copyright of articles printed in ANALYSIS remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.

ANALYSIS

Edited by
C. J. F. WILLIAMS

CONTENTS

- | | |
|---|---------------------|
| The distinction between rhetic and illocutionary acts | STEWART THAU |
| 'Propositional content' | VERA PEETZ |
| A note on modal quantification, ontology and the indenumerably infinite | J. E. WIREDU |
| 'Orders' of sets | A. J. DALE |
| Miss Anscombe on the 'general propositional form' | WILLIAM W. GUSTASON |
| Begging the question | DAVID H. SANFORD |
| Urmson on Russell's incomplete symbols | R. K. PERKINS, JR. |
| Utilitarianism, overall obligatoriness and deontic logic | R. EUGENE BALES |
| Miss Haight on picturing unreality | W. CHARLTON |
| Censorship and participatory democracy: a paradox | W. A. McMULLEN |

ASIL BLACKWELL · BROAD STREET · OXFORD

Forty pence

THE DISTINCTION BETWEEN RHETIC AND ILLOCUTIONARY ACTS

By STEWART THAU

JOHN Austin's distinction between rhetic and illocutionary acts has attracted a particular sort of criticism that does not seem to me to be sensitive to certain important lines of his thought. Of this criticism, the most recent is by John Searle in his article 'Austin on Locutionary and Illocutionary Acts'.¹ An earlier form of this criticism is to be found in an article by L. Jonathan Cohen entitled 'Do Illocutionary Forces Exist?'²

According to Austin, rhetic acts are acts of uttering phemes (words or sentences) 'with a certain "meaning" in the favourite philosophical sense of that word, i.e., with a certain sense and reference'.³ On the other hand, illocutionary acts, though containing rhetic acts, are additionally acts of uttering words with a certain 'illocutionary force'. Now both Searle and Cohen try to show that, at least in some cases, the illocutionary *force* of an utterance cannot be distinguished from the *meaning* of an utterance. They both think this shows that rhetic and illocutionary acts cannot be distinguished. This causes Searle to reject Austin's notion of the rhetic act and Cohen to reject Austin's notion of the illocutionary act. I agree with them that on some interpretations of the rather unclear terms 'illocutionary force' and 'meaning of an utterance' such force cannot be distinguished from meaning and is really one aspect of an utterance's meaning. I do not, however, agree with them that this entails the rejection of either the notion of a rhetic *act* or the notion of an illocutionary *act*. That they should be driven to such rejections is due in part to some confusions that can be traced to Austin himself and in part to a failure to appreciate the possibilities for distinguishing 'illocutionary *acts*' from 'illocutionary *force*'.⁴ Austin himself says a good deal to suggest a plausible and sound way of drawing the distinction between rhetic and illocutionary acts. I will try to bring out this plausible way of drawing the distinction by discussing Searle's form of the criticism.

Why does Searle think that the rhetic and illocutionary acts cannot, at least not always, be distinguished? It is because the *meaning* of an

¹ John R. Searle, 'Austin on Locutionary and Illocutionary Acts', *The Philosophical Review*, Vol. LXXVII, No. 4 (1968).

² L. Jonathan Cohen, 'Do Illocutionary Forces Exist?', *The Philosophical Quarterly*, Vol. 14, No. 55 (1964).

³ John Austin, *How to do Things with Words* (Oxford, 1962), p. 94.

⁴ For a discussion of one of the confusions found in Austin, see my 'Illocutionary Break-downs', *Mind*, Vol. LXXX, April, 1971. Another one of these confusions is mentioned briefly in the last paragraph of the present article.

utterance cannot always be distinguished from its *force*. And why not? Because there are sentences whose meanings determine that they can be correctly used with only one kind of force. For example, whereas 'It's going to charge' can be correctly used with either the force of a warning or a prediction, 'I warn you it's going to charge' can only be correctly used with the force of a warning. If a listener fully understood the meaning of an utterance of the latter sentence, including its explicitly performative prefix 'I warn you', he would understand that it had the force of a warning. So it might seem that the meanings of some utterances contain the forces of those utterances. Searle says:

... but where a certain force is part of the meaning, where the meaning uniquely determines a particular force, there are not two different acts but two labels ("locutionary" and "illocutionary") for the same act (p. 407).

Searle later explains that this is a difficulty not only for utterances of sentences with explicitly performative prefixes but for utterances of every sentence, since every sentence 'has some illocutionary force potential, if only of a very broad kind, built into its meaning' (p. 412). Here he points out that the grammar of a sentence, i.e., whether indicative, interrogative, or imperative, determines the sentence to appropriate usage in some illocutionary acts rather than others. He concludes this line of reasoning by saying:

The point I am making now is that there is no way to abstract a rhetic act in the complete utterance of a sentence which does not abstract an illocutionary act as well, for a rhetic act is *always* an illocutionary act of one kind or another (p. 413, italics mine).

I believe Searle has concluded here much more than he has shown. He might seem to be on the firmest ground when talking of utterances containing explicitly performative prefixes, but even in those cases other examples would have made his conclusions more difficult to draw. For examples, consider the explicitly performative sentences 'I name this ship the Joseph Stalin' and 'I appoint you Justice of the Supreme Court'. Uttering these with a certain appropriate meaning is not sufficient to insure that the speakers have performed the illocutionary acts of naming a ship and appointing a justice. The man naming the ship may not be entitled to name the ship and the man ostensibly being appointed may have already been appointed, or not approved by Congress, etc. Though these speakers have neither *named* nor *appointed*, we can imagine that they have spoken with a certain meaning and that this meaning was quite clear or manifest. According to Austin the rhetic act is the act of speaking 'with a more or less definite sense and a more or less definite reference (which together are equivalent to "meaning")' (Austin, p. 93). So here I have given two examples of utterances that are successful rhetic acts but that are not successful illocutionary

acts. When Searle says that 'there is no way to abstract a rhetic act' from an illocutionary act and that 'a rhetic act is always an illocutionary act', it would seem he is overlooking such cases as I have just described. Perhaps he was thinking only of sentences like 'I promise . . .', 'I apologize . . .', etc., where, if they are uttered with appropriate meaning, very little can frustrate illocutionary success. Or, perhaps, he meant to say that one could not abstract a rhetic act from an *attempt* at an illocutionary act. If he meant this, I believe he is right. However, this would not show that we cannot abstract the one act from the other, for what better way could there be of showing that we can distinguish two types of acts than by showing that the conditions of success for the one (the rhetic) can be met without meeting the conditions of success for the other (the illocutionary)?

Now, even when we switch to such examples as promising and apologizing we can still distinguish rhetic success from illocutionary success. If I say, 'I promise you I will come to your party', or, 'I apologize for stepping on your foot', then, in spite of the fact that I may be speaking with a more or less definite meaning, *you*, the addressee, may not even be present. In such cases I will not have made *you* a promise nor apologized to you. Other things can also go wrong to frustrate illocutionary success if 'success' be taken in a wider sense. If I have not stepped on your foot or you are not giving a party, then my apology or my promise are surely less than fully successful. It was to handle these less than fully successful occurrences in speech, as well as the complete illocutionary failure, that Austin introduced the notion of an *infelicity*. The illocutionary act is the act that is liable to certain types of infelicity or breakdown, and these infelicities can occur even when I have spoken with a more or less definite meaning, i.e., when I have performed a successful rhetic act.

Perhaps it will be thought that I am advocating an interpretation of the distinction between rhetic and illocutionary acts that Searle discusses early in his article and rejects as unsatisfactory. He imagines someone defending the Austinian distinction by saying:

... there is still a distinction between uttering the sentence with (that is, as having) a certain sense and reference on the one hand (the locutionary act) and actually bringing off a *successfully* performed illocutionary act (pp. 408-9).

This imaginary defender is indeed saying something I might say in expressing my own views, especially if 'having a certain sense and reference' is understood in an appropriate way. What according to Searle is wrong with drawing the distinction in this way? He has two objections.

Searle's first objection is:

... it reduces the locutionary-illocutionary distinction to a distinction between trying and succeeding in performing an illocutionary act (pp. 96-7).

To show that Searle's first objection is unsound we have to show that such phrases as 'speaking with a certain meaning' and 'speaking with a more or less definite sense and reference' are not equivalent to '*trying* to perform an illocutionary act', and that, indeed, they denote a certain kind of success in speaking, i.e., rhetic success.

Austin identifies the rhetic act in more ways than by saying it is the act of speaking with a more or less definite meaning. He says:

But the rheme is a unit of *speech*; its typical fault is to be vague or void or obscure, etc. (Austin, p. 98).

Furthermore, in discussing rhetic acts he gives examples that are presumably *rhetic breakdowns*. These examples include such breakdowns as indefinite reference and ambiguity (cf. pp. 96-7). All of this suggests that Austin thought that the ways we fail 'to speak with a more or less definite meaning' are by speaking vaguely, ambiguously, without definite reference, obscurely, etc. Thus, the conditions of rhetic success are that we avoid speaking in any of these ways that frustrate the clarity of our utterance. Now the fact that a speaker is *trying* to perform an illocutionary act hardly insures that he has met these conditions of rhetic success. I may be *trying* to perform the illocutionary act of asking you to meet me at the First National Bank of Syracuse, but since I say only 'Will you meet me at the bank?' it may not be clear just where I want you to meet me. The breakdown may be one of ambiguity (you are not sure in what sense I meant 'bank') or one of indefinite reference (you are not sure which financial institution I meant). In any case, though I was *trying* to perform a certain illocutionary act, I did not succeed even in performing a successful *rhetic act*, i.e., I did not even speak with a more or less definite sense or reference (meaning).

Searle's second objection to the above way of drawing the distinction is less clear. Indeed it does not seem to be an objection at all. He says that this way of drawing the distinction

... is different from the original distinction between an utterance with a particular meaning and an utterance with a particular force (p. 410).

But just what is this 'original distinction'? The terms used in stating it require more explanation than Searle, or even Austin, has given them. What is it for an utterance to '*have a certain meaning*' or for it to '*have a certain force*'? I have just given some reasons for thinking that when Austin talked of an utterance 'with a more or less definite sense and reference' he meant an utterance that, in the context of utterance, manifests or makes clear what the speaker means. But there are at least

ANALYSIS

Edited by C. J. F. Williams

Volume 32

1971-1972

BASIL BLACKWELL • BROAD STREET • OXFORD

INDEX

	PAGE
ATTFIELD, ROBIN	Collective responsibility 31
BALES, R. EUGENE	Utilitarianism, overall obligatoriness and deontic logic 204
BOËR, STEVEN E.	On Searle's analysis of reference 154
BRADY, JAMES B.	Indifference and voluntariness 98
CASTAÑEDA, HECTOR- NERI	On the problem of formulating a coherent act-utilitarianism 118
CASTAÑEDA, HECTOR- NERI	There are command sh-inferences 13
CHARLTON, W.	Miss Haight on picturing unreality 206
CLARK, MICHAEL	Ifs and hooks 33
CLEGG, JERRY S.	Symptoms 90
COOPER, DAVID E.	Searle on intentions and reference 159
DALE, A. J.	'Orders' of sets 192
DAVIS, LAWRENCE H.	They deserve to suffer 136
DUNCAN-JONES, AUSTIN	Dissent from precepts 1
FINNIS, JOHN	The restoration of retribution 131
GILBERT, MARGARET	The abilities of prescriptivism 141
GUSTASON, WILLIAM	Miss Anscombe on the 'general proposi- tional form' 195
HAIGHT, M. R.	Picturing unreality 65
HAKSAR, VINIT	Rawls' theory of justice 149
HARRISON, JONATHAN	The letter 86
HILL, CHRISTOPHER S.	'gavagai' 68
HOFFMAN, ROBERT	On begging the question at any time 51
HUMBERSTONE, I. L.	Two sorts of 'ought's 8
JOHNSEN, BREDO C.	Black and the inductive justification of induction 110

JONES, O. R.	Knowing and guessing—by examples	19
JONES, O. R.	Truth and predication	104
KORDIG, CARL. R.	Moral weakness and self-reference	11
LINSKY, LEONARD <i>and</i>		
SCHUMM, GEORGE F.	Frege's way out: a footnote	5
LYCAN, WILLIAM G. <i>and</i>		
OLDENQUIST, ANDREW	Can the generalization argument be reinstated?	76
McMULLEN, W. A.	Censorship and participatory democracy: a paradox	207
MANNISON, D. S.	Matthews on enjoying digging	173
MARTIN, MICHAEL	Confirmation and explanation	167
MORICK, HAROLD	Intentionality, intensionality and the psychological	39
MYRO, GEORGE	A note on strict implication and entailment	55
NICKEL, JAMES W.	Discrimination and morally relevant characteristics	113
OLDENQUIST, ANDREW	(<i>See</i> Lycan, William G.)	76
PEETZ, VERA	'Propositional content'	183
PERKINS, JR., R. K.	On Russell's alleged confusion of sense and reference	45
PERKINS, JR., R. K.	Urmson on Russell's incomplete symbols	200
RADFORD, COLIN	On sticking to what I don't believe to be the case	170
SANFORD, DAVID H.	Begging the question	197
SAYWARD, CHARLES	True propositions: a reply to C. J. F. Williams	101
SCHUMM, GEORGE F.	(<i>See</i> Linsky, Leonard)	5
SIMPSON, R. L.	Stoics, sleepers and stones	164
SINGER, PETER	Moral experts	115
SLOMAN, AARON	New bodies for sick persons: personal identity without physical continuity	52

SMART, BRIAN	How to reidentify the ship of Theseus	145
SPANGLER, G. A.	Mannison on doing something on purpose but not intentionally	100
THAU, STEWART	The distinction between rhetic and illocutionary acts	177
WALLACE, G.	Cargile on Utilitarianism	126
WEST, DAVID	Phillips, Mounce and repaying what we borrow	24
WESTPHAL, FRED	Utilitarianism and "conjunctive acts": a reply to Professor Castañeda	82
WHITE, ALAN R.	Meaning and implication	26
WILLIAMS, C. J. F.	Truth: a composite rejoinder	57
WIREDU, J. E.	A note on modal quantification, ontology and the indenumerably infinite	187
ZELLNER, HAROLD N.	Utilitarianism and derived obligation	124

two other interpretations for that phrase, neither of which would, I think, be good interpretations of Austin's intention. (1) It might mean that a speaker has uttered a sentence that has a meaning in a given language. On this interpretation, if a speaker had uttered, 'The slithy toves did gyre', he would not have spoken 'with a more or less definite meaning'. (2) It might mean, on the other hand, that a speaker has uttered a sentence and *meant something by it*. If taken in this latter sense, the fact that someone had spoken 'with a more or less definite meaning' would not imply that a speaker had spoken clearly or even used meaningful sentences of a language. I may have *meant* (in this second sense) the First National Bank even though my reference to it was indefinite or ambiguous and I may have meant that the door was closed by saying 'The slithy toves did gyre'. Let us use suffixes to keep these senses of 'meaning' distinct. When I use the word in the sense described in (1) above, I will suffix '-*p*' to it. I do this because speaking with a meaning-*p*, i.e., uttering a meaningful sentence of a language, is a condition of performing a *phatic* act. On the other hand, when I use the word in the sense described in (2) above, I will suffix '-*m*' to it. It is of meaning-*m*, primarily, that H. P. Grice was trying to give us an analysis. His analysis was in terms of intentions.¹

Now, what I have been arguing is that Austin was using the phrase 'speaking with a more or less definite meaning' neither in the sense of 'speaking with a meaning-*p*' nor in the sense of 'speaking with a meaning-*m*'. It is possible for us to speak with a "meaning" in both of these senses and still not to have performed a rhetic act. When we speak ambiguously, for example, we utter sentences that have a meaning-*p*; indeed they usually have too many meanings-*p*. Also, when we speak ambiguously, we often *mean-m* something quite definite—we just fail to make what we mean-*m* clear. To proliferate terminology, let us suffix '-*r*' to 'meaning' in Austin's sense, since to speak 'with a meaning' in that sense implies rhetic success. When we say of a speaker, then, that he has spoken with a meaning-*r*, we are not just saying that he uttered a sentence with a meaning-*p* nor are we just saying that he uttered a sentence and *meant-m* something by it; we are saying that the sentence he uttered, in the context of utterance, *appropriately manifests* what he *meant-m* by the sentence.

Now the question at issue can be formulated as follows: Is there a distinction between an utterance with a meaning-*r* and an utterance with an illocutionary force? Before we can answer this we must make another digression and ask what it is for an utterance to 'have an illocutionary force'. Here we might mean any one of three things: (1) We might mean that a speaker uttered a sentence and *meant-m* it as a given illocutionary act. (2) We might mean not only that he *meant-m* it

¹ H. P. Grice, 'Meaning', *The Philosophical Review*, Vol. LXVII (1957).

as a given illocutionary act but also that he succeeded in making clear which illocutionary act he meant-*m* it as. (3) We might also mean that, in addition to making clear what illocutionary act he meant-*m* it as, he uttered his sentence in circumstances that constitute his utterance a successful or felicitous illocutionary act. I have found it useful to say that an utterance has a given illocutionary force if the utterance (including its context) makes clear just what illocutionary act the speaker means-*m* to be performing. This is the second sense listed above. If we adopt this stipulation, it becomes possible to say of certain utterances that they have the force of, say, naming or appointing, though they are not successful acts of naming or appointing.

To repeat the question once more: Is there a distinction between an utterance with a meaning-*r* and an utterance with a force? I think not. Speaking with an illocutionary force is one aspect of speaking with a meaning-*r*, i.e., one aspect of performing a rhetic act. Rhetic acts are the acts of making clear or *appropriately manifesting* what we mean-*m*. We fail to perform these acts when we fail to make clear what illocutionary acts we mean-*m* to perform as much as when we fail to make clear what we are referring to or in what sense we are using certain words. But though we cannot distinguish speaking with a certain force from one aspect of speaking with a meaning-*r*, we *can* distinguish performing a *successful* or *felicitous illocutionary act* from speaking with a meaning-*r*, that is, we can distinguish illocutionary acts from rhetic acts.

Searle's second objection to the above way of drawing the rhetic-illocutionary distinction is that it is different from 'the original distinction between an utterance with a particular meaning and an utterance with a particular illocutionary force'. If it is different, then at least it has its own virtues, for the 'original distinction' is one that will not hold up in the way intended. But I think that the vagueness in the notion of an utterance's 'having an illocutionary force' should caution us not to think that the original distinction is completely clear.

Thus, it seems to me that Searle has not shown that there is any need to reject Austin's notion of a rhetic act. I think, however, that he is reacting to some of the confusion in Austin's own thoughts concerning rhetic acts. Austin seems to have been in at least two minds with respect to these acts. On the one hand, they are acts that we fail to perform when we speak ambiguously, with indefinite reference, vaguely, obscurely, etc. This is the aspect of Austin's intention that I have emphasized. On the other hand, Austin did say that with the constative utterance 'we abstract from the illocutionary . . . aspects of the speech act and we concentrate on the locutionary. . . .' (Austin, pp. 144-5). This suggests that Austin was interested in what is expressed in the subordinate clauses of explicitly performative formulas (e.g., 'that your barn is on fire' of 'I warn you that your barn is on fire') and that he *sometimes*

thought of what was expressed in those subordinate clauses as the upshot of the locutionary (or rhetic) act. Searle has sought to draw out this intention of Austin's in his own notion of a *propositional act*. I have said nothing against Searle's endeavours in that direction. What I have maintained is that 'speaking with a more or less definite meaning' may be considered a kind of success we attain in speaking and that it is not identical with either (1) *trying* to perform an illocutionary act or (2) successfully performing an illocutionary act. Since Austin defines a rhetic act as the act of speaking with a more or less definite meaning, I think we can say that Searle has not shown that we ought to reject Austin's notion of a rhetic act. The criticism that Austin deserves is that he thought he could handle too many concerns with this one notion, i.e., the rhetic act cannot be both the act of speaking with a more or less definite meaning *and* a propositional act.

Syracuse University

'PROPOSITIONAL CONTENT'

By VERA PEETZ

CONSIDER the following illocutionary acts:

- | | | |
|---|---|--|
| A | { | I assert that the cat is on the mat |
| | { | I guess that you are holding the seven of diamonds |
| | { | I predict that Pegasus will win |
| B | { | I promise to give you a pony |
| | { | I order you to feed the ducks |
| | { | I warn you to keep away from the bull. |

Professor John Searle would consider that all these illocutionary acts are exemplifications of the formula $F(x)$, where 'F' stands for the illocutionary force of the utterance and 'x' stands for the propositional content. 'Propositional content' is an expression used by Searle¹ to denote what is common to, for example, 'I assert that John Smith shut the door', 'I, John Smith, promise to shut the door', 'John Smith, shut the door!', 'Did John Smith shut the door?', and so on, namely the proposition 'John Smith shut the door'. However, I think that there are important differences between the illocutionary acts in group A and the illocutionary acts in group B, which suggest that Searle's analysis of an illocutionary act into propositional content plus illocutionary force is not quite correct.

¹ See his book, *Speech Acts*, Cambridge University Press, 1969.

In this paper, I shall use 'the content of the illocutionary act' to refer to what follows (in the sense of the import of the words used not the actual words themselves) the illocutionary-force-bearing verb in the above illocutionary acts; that is, what is asserted, what is predicted, what is promised, what is ordered, and so on. Now illocutionary acts of the type found in group B can be equally well performed with their contents in the same form as the contents of the illocutionary acts in group A: 'I promise to give you a pony' can be equally well performed by 'I promise that I will give you a pony', 'I order you to feed the ducks' can be equally well performed by 'I order that you feed the ducks' and 'I warn you to keep away from the bull' can be equally well performed by 'I warn that you (should) keep away from the bull'. The converse sometimes holds, sometimes does not hold; that is, illocutionary acts of the type found in group A can sometimes be performed with their contents in the same form as the contents of the illocutionary acts in group B, and sometimes they cannot. 'I assert that the cat is on the mat' can be performed equally well by 'I assert the cat to be on the mat', but 'I predict that Pegasus will win' cannot be performed equally well by 'I predict Pegasus to win'.

In group A the illocutionary acts are acts of asserting *that* such-and-such, guessing *that* such-and-such and predicting *that* such-and-such. These acts can be symbolised by 'I R *p*', where 'R' represents an illocutionary-force-bearing verb (bearing its illocutionary force) and '*p*' represents a proposition. Even when acts such as those in section A can be expressed in the form, for example, 'I assert the cat to be on the mat', one is still asserting something to be the case, that is, one is asserting a proposition. In group B, however, the illocutionary acts are acts of promising *to perform* some act, ordering someone *to perform* some act and warning someone *to perform* some act. It is true, as I have said above, that these acts can be uttered with their contents in apparently propositional form. But in spite of this, I do not think that these acts can be symbolised by 'I R *p*', where '*p*' stands for a proposition, as can the acts in group A. Whereas a proposition can be asserted, guessed, or predicted, it cannot be promised, ordered or warned (in the way in which 'warn' is used in group B). I promise to perform some act; I do not promise a proposition or to perform a proposition. What I could be said to be doing is promising to make a proposition true, but this is still to promise an act, not a proposition.

Now Searle analyses 'I promise to come' as the proposition 'I will come' plus the illocutionary force 'I promise'; but, in view of what I have said above, the analysis should really be the proposition 'I will come' plus 'I promise to make this proposition true'. However, now we have, not a proposition plus the illocutionary force with which it is uttered, but a proposition plus an illocutionary act which in its turn

requires analysis. So the analysis of 'I promise to come' will be proposition₁ 'I will come' plus proposition₂ 'I will make proposition₁ true' plus 'I promise to make proposition₂ true'. Again, the analysis involves an illocutionary act which must be analysed, and this process will be repeated *ad infinitum*. It must be concluded, therefore, that the sort of analysis which Searle gives leads to an infinite regress and is therefore unsatisfactory.

Let us look at the illocutionary acts in section B in their so-called propositional form, that is, in the form 'I promise that I will give you a pony', and so on, and let us compare these illocutionary acts with the illocutionary acts in section A. In certain contexts, it is possible to utter 'The cat is on the mat', 'You are holding the seven of diamonds' and 'Pegasus will win' (without explicit indication of the respective illocutionary forces attaching to them) and be taken to be performing, respectively, an act of asserting, an act of guessing and an act of predicting. Similarly, in certain other contexts, 'I will give you a pony', 'You feed the ducks' and 'You (should) keep away from the bull' can be uttered (also without explicit indication of the respective illocutionary forces attaching to them) and be taken respectively as an act of promising, an act of ordering and an act of warning. This suggests that 'I will' in 'I will give you a pony' has something of the 'I will' of willing or intending rather than being a means of producing the future tense of 'give'; that the verb in 'You feed the ducks' is imperative rather than indicative, and that there is also something of the imperative about the verb in 'You (should) keep away from the bull'. For these reasons, I would suggest that it is wrong to label 'I will give you a pony', 'You feed the ducks' and 'You (should) keep away from the bull' as propositions. Just as there are two senses of 'assertion' or 'guess' or 'prediction': the illocutionary act of asserting or guessing or predicting, on the one hand, and what is asserted or what is guessed or what is predicted, on the other, so there are two senses of 'promise' or 'order' or 'warning': the illocutionary act of promising or ordering or warning, on the one hand, and what is promised or what is ordered or what is warned, on the other. In the case of assertions, guesses and predictions (in the sense of what is asserted, what is guessed and what is predicted) these can always be expressed in the form of propositions which have a truth value; but in the case of promises, orders and warnings (in the sense of what is promised, what is ordered, what is warned), because the verbs are not indicative (or certainly not purely indicative) then, although they can assume what looks like propositional form, they are not really propositions and do not have truth values. Hence illocutionary acts of the type found in group A are acts of asserting that such-and-such is the case, predicting that such-and-such will be the case, guessing that such-and-such was/is/will be the case, and so on; or, to put it another way, these acts are acts

where a proposition is asserted, guessed, predicted, and so on. Illocutionary acts of the type found in section B are acts of promising to perform some act, ordering someone to perform some act, warning someone to perform some act, and so on; or, to put it another way, these acts are acts in which a promise is promised, an order is ordered, a warning is warned (given).

Hence Searle's analysis of an illocutionary act as proposition plus illocutionary force will apply only to the type of illocutionary act found in section A. For illocutionary acts of the type found in section B, there is no single comparable analysis which can be given of them (for a promise—in the sense of what is promised—is not the same kind of thing as an order—in the sense of what is ordered, and so on). However, the analysis could be made more general (so that it would cover not only the type of illocutionary act found in section B, but also the type found in section A): an illocutionary act could be analysed as illocutionary force plus the content appropriate to that illocutionary force. We could symbolise this as $F(x)$, where 'F', as before, stands for illocutionary force, but where now 'x' stands not for propositional content but content appropriate to that particular kind of illocutionary force (that is 'x' will stand for a proposition, a promise, an order, a warning, and so on, depending on the illocutionary force for which 'F' stands).

Hence, Searle's concept of propositional content is, I think, misleading. If he wants a name for what is common to, for example, 'I assert that John Smith shut the door', 'I, John Smith, promise to shut the door', 'John Smith, shut the door', 'Did John Smith shut the door?', and so on, I suggest that 'corresponding proposition' is a better name; it is the corresponding proposition of these various illocutionary acts in the sense that it is the proposition which corresponds to the content of each of them.

University of Nottingham

A NOTE ON MODAL QUANTIFICATION, ONTOLOGY AND THE INDENUMERABLY INFINITE

By J. E. WIREDU

QUINE has repeatedly insisted that quantification in modal logic has queer ontological consequences. 'It leads us' he says 'to hold that there are no concrete objects (men, planets, etc.), but rather that there are only, corresponding to each supposed concrete object, a multitude of distinguishable entities (perhaps, "individual concepts", in Church's phrase).' ('The Problem of Interpreting Modal Logic', *The Journal of Symbolic Logic* Vol. 12, 1947. All page references are to the reprint in Irving M. Copi and James A. Gould: *Contemporary Readings in Logical Theory*, New York, Macmillan 1967.) I start by examining Quine's argument for this contention.

To begin with, Quine elicits from the thinking of modal quantificationists the following principle (numbered (ii) in his text):

An existential quantification holds if there is a constant whose substitution for the variable of quantification would render the matrix true (*ibid.*, p. 271).

In Quine's own view this can only be a partial criterion of quantificational interpretation, but he adopts it hypothetically to 'see how it fares'. Now comes the 'proof' of the untoward ontological consequences:

... let us use C (for 'congruence') to express the relation which Venus, the Evening Star and the Morning Star, e.g., bear to themselves and, according to empirical evidence, to one another. ... Then

Morning Star C Evening Star . \square (Morning Star C Morning Star).

Therefore, according to (ii),

(1) $(\exists x) (x \text{ C Evening Star} . \square (x \text{ C Morning Star}))$.

But also

Evening Star C Evening Star . $\sim \square$ (Evening Star C Morning Star),

so that, by (ii),

(2) $(\exists x) (x \text{ C Evening Star} . \sim \square (x \text{ C Morning Star}))$.

Since the matrix quantified in (1) and the matrix quantified in (2) are mutual contraries, the x whose existence is affirmed in (1) and the x whose existence is affirmed in (2) are two objects; so there must be at least two objects x such that $x \text{ C Evening Star}$. (Quine: *ibid.*, p. 272.)

A good number of logicians have offered replies to this argument of Quine's, notably, F. B. Fitch, 'The Problem of the Morning Star and the Evening Star', *Philosophy of Science*, Vol. 16 (1949), reprinted by Copi and Gould, op. cit. pp. 273ff; A. F. Smullyan, 'Modality and Description', *The*

Journal of Symbolic Logic, Vol. 13 (1948), and Ruth Barcan Marcus, 'Modalities and Intensional Languages', *Synthese*, Vol. 27 (1962), also reprinted by Copi and Gould, op. cit. pp. 278ff. But, to my knowledge, none has made the point that in this argument Quine simply begs the main question. Given (ii), (1) and (2) merely lead to the inference that there must be at least two constants such that their substitution in ' α C Evening Star' renders the matrix true. It may well be, of course, that (ii) is an inadequate principle for interpreting the existential quantifier, but in the argument in question it happens to be the principle being tested. Hence it was not open to Quine to invest the existential quantifications in (1) and (2) with his own 'object oriented' interpretation. So the argument collapses.

Nevertheless, the issue of the ontologically neutral, versus the 'object oriented', interpretation of quantifiers is important enough in itself to merit independent investigation. It turns out, in fact, that it is a crucial bone of contention between Quine and modal quantificationists of the persuasion of Ruth Barcan Marcus. (See the reply of Marcus, *loc. cit.*, and Quine's rejoinder, 'Reply to Professor Marcus', *Synthese*, Vol. 27, 1962, also reprinted by Copi and Gould, op. cit. pp. 293ff.) In her reply, Marcus advances just the type of interpretation of quantifiers which Quine formulates in (ii). To this Quine urges two objections in his rejoinder. The second objection can, I think, be disposed of fairly expeditiously, and so I notice it first. According to Quine, to adopt the neutral interpretation is to abstract from reference altogether: 'Quantification ordinarily so called is purely and simply the logical idiom of objective reference' (Quine, in Copi and Gould, op. cit. p. 298). The import of this remark is, presumably, that it is essential for the purposes of logic—i.e., for the systematic study of the logical structure of inference or, what in my opinion comes to the same thing, for the systematic investigation of logical truths—that in phrases such as 'there is an α such that . . . ' and 'for all α . . . ' the variable α should be taken as a means for referring to objects. This appears to be false. Consider the following argument:

Anarchism is utopian
 All things utopian are subversive
 \therefore Anarchism is subversive.

There is no reason why this argument should not be symbolised as follows (using ' a ' for 'anarchism', ' U ' for 'utopian' and ' S ' for 'subversive'):

Ua
 $(x)(Ux \supset Sx)$
 $\therefore Sa$.

The argument is clearly formally valid. Correlatively, one might point out that 'if $[Ua \cdot (x)(Ux \supset Sx)]$ then Sa ' is a quantificational truth. We now observe that generalization on Ua , for instance, yields $(\forall x) Ux$; whereupon, on Quine's 'real' interpretation of quantifiers, one is implausibly constrained to admit that anarchism is an object.

Quine is, of course, not the one to be nonplussed by an argument such as this. In fact in *Word and Object* (Cambridge, The M.I.T. Press, 1960), he pointed out that the ontic content of a piece of discourse may not become clear until it is reformulated in 'the canonical notation of quantification'. (See, e.g., §25 and chap. VII.) Only those terms which can survive 'the canonical expedient of reparsing singular terms' are to be admitted to name objects. In the present case, 'anarchism' might well succumb to 'canonical' reparsing. It might be claimed that the statements about anarchism given above might be reformulated in such a way as to be seen to be about anarchists and their dispositions—dispositions, in turn, being presumably susceptible of a behavioural construal. That might be granted. But it is beside the point, which is this. In order to appreciate the formal validity of the argument featuring the statements about anarchism or to see the logical truth of the corresponding conditional statement it is clearly not relevant to embark upon the kind of analytical experiment hinted at. Any linguistic analysis and consequent rephrasing of a given argument is *logically* relevant only to the extent that it enables us to put it into logical form in some domain of logic for formal appraisal. Since in the case of the presented example the argument is immediately amenable to quantificational treatment, any interest attaching to an attempt to reformulate it further must be declared extra-logical, i.e., beyond the sole aim of subjecting it to quantificational evaluation. Accordingly we stand by our objection that if the logical idiom of quantification of itself carried objectual reference, then the cited example of an argument would commit us to making an object of anarchism; which, surely, must be a *reductio ad absurdum* of the 'object oriented' interpretation.

In his 'Metaphysics in Logic' (in Flew (ed.) *Essays in Conceptual Analysis*, London, Macmillan, 1956), G. J. Warnock made a similar criticism of the automatically objectual construal of quantifiers, taking 'Valhalla is mythological' as his example for existential generalization. (I may say that I borrow the word 'objectual' from Quine's *Philosophy of Logic*, Prentice Hall, 1970, pp. 93, 94.) It is interesting to note that Bertrand Russell in a retort to Warnock insisted with some heat that the existential quantifier is ontologically neutral. (See Russell: 'Logic and Ontology' in *My Philosophical Development*, Allen and Unwin, 1959, p. 234.) In doing so Russell was, in fact, only reaffirming with maximum explicitness the interpretation which he and Whitehead gave none too clearly in *Principia Mathematica* in the original introduction. In view of

the historical priority of the neutral interpretation, it is not immediately clear what Quine means when he calls his 'object oriented' interpretation the 'ordinary' one.

Quine's first objection against the ontologically neutral interpretation of quantifiers such as is advocated by Marcus is very much more difficult to handle. With respect to that interpretation, he argues that:

... it deviates from the *ordinary* interpretation of quantification in ways that can matter. For one thing, there is a question of unspecifiable objects. Thus take the real numbers. On the classical theory, at any rate, they are indenumerable, whereas the expressions, simple and complex, available to us in any given language are denumerable. There are, therefore, among the real numbers, infinitely many none of which can be *separately* specified by any expressions, simple or complex. Consequently an existential quantification can come out true when construed in the *ordinary* sense, thanks to the existence of appropriate real numbers and yet be false when construed in Prof. Marcus's sense, if by chance those appropriate real numbers all happen to be severally *unspecifiable*. (Quine in Copi and Gould, op. cit. p. 298. My italics.)

The point of special interest for us here lies in the fact that Quine apparently supposes that there can be a *discrimination* among real numbers of some that are separately specifiable and others that are not. But, of course, no *particular* real number can defy specification, for if it is *given*, then it is to that extent captured, whereas if it is 'ungiven' the question of *its* specifiability or unspecifiability does not arise.

We come up, then, against the following paradox: Although there are more real numbers than there are expressions—interpreting the word 'more' in the familiar technical sense—it is, nevertheless, plainly false of any one real number that it is not separately specifiable. This is only an apparent paradox, however. Consider the claim 'There is at least one real number that is not specifiable'. What prevents us from specifying *that* number? As soon as this question is considered, it becomes evident that the fact that the real numbers are indenumerably infinite is completely inoperative in any particular case, for any one real number is *one* and not infinite. It is therefore the case that for any particular real number an expression can be found with still a denumerably infinite stock of expressions to spare! We can see, then, that no one real number is logically destined to be intrinsically anonymous. It is clear in consequence that Quine's suggestion that the real numbers appropriate to some given existential quantification might by chance 'all happen to be severally unspecifiable' is radically unrealisable. Thus the ontologically neutral interpretation stands unprejudiced.

In 1953 Barkley Rosser in his *Logic for Mathematicians*, New York, McGraw-Hill Company, exploited the idea of so called unspecifiable objects to try to establish that not every class has a determining condition. He proceeded as follows:

To show that not every class has a determining condition, we give the following proof, known as Skolem's paradox. . . . For each real number, we can determine at least one class of real numbers; for instance, the class of all smaller real numbers. . . . Thus, the set of all classes of real numbers is not denumerable, since the set of all real numbers is not denumerable. Then the set of all classes whatsoever is certainly not denumerable. So, if every class has a determining condition, the set of conditions must be nondenumerable. As every condition is a statement, the set of statements must accordingly also be nondenumerable. However, the set of statements is denumerable . . . (p. 199)

Interestingly, Rosser himself, in a bid to allay any excessive alarm that might be occasioned by this paradox, remarked:

. . . it really doesn't matter that there should be classes with no determining conditions because no one will ever exhibit such a class. As soon as one exhibits some particular class, one can find a condition which determines that class, for if no better condition is available, one can take the condition of being a member of that class. Since we are dealing with an explicit class, this will give an explicit condition for that class (*ibid.*, p. 200).

In this passage, Rosser came as close as can be imagined to recognising that one cannot speak of a class that is not specifiable. What, after all, is a non-explicit class? Is it intelligible to suppose that there are some classes which are eclipsed in an original shadow of obscurity? Presumably, even such classes can be comprehended under the putative condition: 'The class of all classes that are eclipsed in an original shadow of obscurity'. From which one might, perhaps, derive the maxim, 'For explicit classes, explicit conditions shall be given; for non-explicit classes, non-explicit conditions'.

Finally, it should be observed that it is inaccurate to conceive of language as a set of statements. Language is better thought of as a set of rules for generating statements or, more generally, expressions. Expressibility depends not on what expressions exist but on what expressions *can* be constructed. Actually, Cantor's 'diagonal' proof that the set of real numbers is larger than the set of rationals can be pressed into service in this context. The diagonal argument shows that in any attempt to make a one-one correlation of the rationals with the reals, there will be at least one real number, call it r° , that eludes the enumeration. The crucial consideration here is that r° , the diagonal excess, is itself, nevertheless, specified in a language; and it is obvious that any diagonal excess in any attempted enumeration can be similarly designated. This is simply a reflection of the fact that, after all, it is in a language that the real number series is defined in the first place.¹

¹ I wish to thank my colleague Derek Gjertsen for helpful criticisms of an earlier draft of this paper.

'ORDERS' OF SETS

By A. J. DALE

IN this paper I wish to expose a mistake first pointed out, as far as I know, by Ramsey,¹ which now reappears in modern guise in a series of papers by Tucker.²

It was pointed out by Ramsey in a discussion of Russell's axiom of reducibility that the property of being elementary or non-elementary belongs, not to a proposition as Russell claimed, but more properly to an expression of that proposition. That is, one and the same proposition may be expressed by an elementary expression and a non-elementary expression. Thus Russell's hierarchy of orders stratifies propositional symbols rather than propositions.

In Tucker's papers we find the expressions 'first-order' and 'second-order' applied to sets, and he uses the difference between first-order sets and second-order sets to explain how Cantor's diagonal procedure works.

In the second order pattern of argument for indenumerability, we start with the denumerable set of positive integers, and take any series of subsets which can be put into one-one correlation with the given set. We then define N as the subset of all those integers which are not members of the subset with which they are correlated, and since N is a subset which cannot be correlated with an integer, we have produced the required indenumerable case. N is not of the same order as the other subsets since it is a condition of membership of N to be already correlated with a subset, and N is accordingly a subset of second order which depends for its members on subsets of first order. Since N is parasitic upon the establishment of the first order correlation, it cannot also be part of it.³

Now I claim that the notion of 'order' applied here properly belongs to the *definition* of the set rather than the set itself. We have in the case of sets, as Ramsey had not in the case of propositions, a clear criterion of identity: sets are identical if they have the same members. It is possible, and I shall give examples, for a set to have a first-order definition and a second-order definition, thus showing that talk of orders is certainly mistaken as applied to sets. I shall also show that Tucker is wrong in claiming that 'since N is parasitic upon the establishment of the first-order correlation, it cannot be part of it'. It is certainly true that N is not one of the sets of the correlation but not for the reason he gives. It has nothing to do with its being parasitic.

I shall consider not the set of all sets of natural numbers but the set

¹ F. Ramsey, *The Foundations of Mathematics*, p. 34.

² J. Tucker, 'The Formalisation of Set Theory', *Mind*, 1963, pp. 500-18; 'Constructivity, Consistency and Natural Languages', *Proceedings of the Aristotelian Society*, 1966-1967, pp. 145-68; 'Rules Automata and Mathematics', *Proceedings of the Aristotelian Society*, 1969-1970, pp. 161-79.

³ 'Rules Automata and Mathematics', p. 173.

of all *finite*¹ sets of natural numbers. I do so because this set is equivalent to the set of natural numbers and a correspondence can be established between each natural number and each finite set of natural numbers, thus simplifying the construction of "second-order" sets. The correspondence can be established by means of arranging the sets in a sequence: A precedes B in the sequence if the sum of the members of A is less than the sum of the members of B; if their sums are equal A precedes B if the least member of A is less than the least member of B, or if these are equal, if the next to least of A is less than the next to least member of B, and so on. It is clear that before any set in this sequence there will only be a finite number of sets at most and that all finite sets of natural numbers can be reached in this way after only a finite number of sets. Hence, a correspondence has been set up, the first few terms of which are:

- 1 \leftrightarrow {1}
- 2 \leftrightarrow {2}
- 3 \leftrightarrow {1, 2}
- 4 \leftrightarrow {3}
- 5 \leftrightarrow {1, 3}
- 6 \leftrightarrow {4}
- 7 \leftrightarrow {1, 4}
- 8 \leftrightarrow {2, 3}
- 9 \leftrightarrow {5}
- 10 \leftrightarrow {1, 2, 3}
- 11 \leftrightarrow {1, 5}
- 12 \leftrightarrow {2, 4}
- 13 \leftrightarrow {6}
- 14 \leftrightarrow {1, 2, 4}
- 15 \leftrightarrow {2, 5}
- 16 \leftrightarrow {3, 4}
- 17 \leftrightarrow {7}

In this correspondence it is clear that the set of all natural numbers which do not belong to their corresponding sets (a second-order set in Tucker's usage) is the set of all numbers greater than 2 (a first-order set). So that although the set has been defined by means of a second-order expression it does not entail that the set is of "second-order". The differences between "second-order" and "first-order" as applied directly to sets is seen to be unreal.

The correspondence also illustrates a much greater objection to Tucker's argument. On his own account, the "second-order" set N cannot be part of the correlation. On exactly similar grounds he would have to admit that the same reasoning applied to the set of all positive

¹ In the case I am discussing, it is the set of all *non-empty* finite sub-sets of the set of all natural numbers.

integers that were members of their corresponding sub-sets. For this set too is a "second-order" set parasitic for its members on "first-order" sets. In the above correspondence between the set of positive integers and the set of all finite sub-sets of this set the set of all positive integers which are not members of their corresponding sub-sets is the set of all positive integers greater than 2 and the set of all positive integers which are members of their corresponding sub-set is the set (1, 2). Since there is no positive integer corresponding to the set of all integers greater than 2 (the set being infinite) the question of whether the n which corresponds to it is a member of the set does not arise. But the question may be asked of the other "second-order" set, the set of all positive integers which do belong to their corresponding sub-sets and, in this case, answered negatively, for 3 (which corresponds to the set (1, 2)) does not belong to (1, 2). By interchanging in the above correspondence the set (1, 2) and (1) we find that the set of all integers which do belong to their corresponding sub-sets—the set (1, 2) again—is correlated with an integer, 1, which *does* belong to its corresponding set.

Thus we see by the above examples that 'first-order' and 'second-order' are applicable to definitions and not to sets, and also that the fact that a set has a second-order definition does not imply that it must be different from the first-order sets of the correlation.

University of Hull

MISS ANSCOMBE ON THE 'GENERAL PROPOSITIONAL FORM'

By WILLIAM W. GUSTASON

IN chapter ten of her *Introduction to Wittgenstein's Tractatus*,¹ Miss Anscombe ably expounds Wittgenstein's concept of the general form of proposition, and in particular his claim that any proposition must occur as a term in the formal series of propositions generated by repeated applications of the operation of joint denial on the set of elementary propositions. Such a series, as she shows, forms a simple progression, thus allowing us to specify in principle any proposition by a description of the form: 'the result of the n^{th} application of $N(\xi)$ to such-and-such ordering of elementary propositions'. It is this series and the functional completeness of the joint denial operation which provide support for one of the most fundamental claims of the *Tractatus*—that 'all propositions are results of truth-operations on elementary propositions' (5.3). Anything sanctioned by the picture theory as a legitimate, "sayable" proposition must occur as a term in the formal series. But she concludes her discussion with the following objection:

However, the theory of the *Tractatus*, promising though it looked at the time, has been clearly and cogently refuted. . . . If all truths of logic are tautological truth-functions of elementary propositions, then there is in principle a decision procedure for them all. But it was proved by Church in the 1930's that multiple quantification theory has no decision procedure; that is, that there cannot be a method by which one could settle, concerning any well-formed formula of that theory, whether it was a theorem or not (p. 137).

Since propositions of quantification theory do not admit of a decision procedure whereas those generated by applications of joint denial clearly do, it would thus follow that there are propositions which cannot occur in Wittgenstein's series.

That Miss Anscombe is mistaken here can be seen, I think, by noting an ambiguity in Wittgenstein's use of 'proposition'. In the 6.1's where he discusses logical truth, he speaks of the "propositions" of logic, of "tautologies". The use of the plural clearly suggests he would count, say, ' $p \vee (\sim p \vee q)$ ' and ' $((p \supset q) \& \sim p) \supset q$ ' as different propositions or tautologies. However, it is claimed earlier that 'a proposition is the expression of its truth-conditions' (4.431), and that understanding a proposition consists in knowing what is the case if it is true (4.024). This well-known doctrine requires that if p and q have different truth-conditions, then they are different propositions, or alternatively:

- i. If p and q are the same proposition, then p and q are true under the same conditions.

¹ Harper and Row, New York (1966), third edition.

Moreover, we are told at 5.141: 'If p follows from q and q from p , then they are one and the same proposition'. And 'following from' is explained at 5.12: '... the truth of a proposition " p " follows from the truth of another proposition " q " if all the truth-grounds of the latter are truth-grounds of the former'. We are thus assured that:

- ii. If p and q are true under the same conditions, then p and q are the same proposition.

From (i) and (ii) it thus follows that logical equivalence is Wittgenstein's criterion of identity for propositions; thus, while ' $\sim(p \vee q)$ ' and ' $\sim p \ \& \ \sim q$ ' are distinct propositional *signs*, they are the same *proposition*.

It should now be apparent that *this* notion of 'proposition'—hereafter 'proposition₁'—is at variance with that of the 6.1's inasmuch as any pair of tautologies are logically equivalent. The latter notion—hereafter 'proposition₂'—appears to coincide with that of a propositional sign. To the extent, then, that tautologies are propositions at all, ' $p \vee (\sim p \vee q)$ ' and ' $((p \supset q) \ \& \ p) \supset q$ ' are different propositions₂ though the same proposition₁.¹ It now becomes clear that, first, Wittgenstein's thesis that all propositions are truth-functions of elementary propositions is a thesis about propositions₁, and second, Church's result that the wffs of quantification theory admit of no decision procedure concerns propositions₂. As to the first point, Wittgenstein's claim is surely that every distinct, "sayable" item of information in natural languages—in effect every distinct set of truth-conditions—can find its expression by a term of the formal series; thus it will contain all truth-functions of one argument, two arguments and so forth. This is clearly a claim about propositions₁; it does not matter if the series can generate all propositions₂ which are the *same* proposition₁, but only that at least one such proposition₂ be produced. On the other hand, Church's result concerns wffs of quantification theory, hence propositions₂ or propositional signs. And the theorems of that theory—those propositions₂ which are logically true—are all the very *same* proposition₁. Thus it is not pertinent whether any of them occur in Wittgenstein's series or not; what matters is that the series contain at least one proposition₂ which may be truly described as the same proposition₁ as they. That there is such a proposition₂ is evident from Miss Anscombe's own exposition of how the series may be generated, for the series obviously contains ' $N[N(\dot{p}, N(\dot{p}, \dot{p})), N(\dot{p}, N(\dot{p}, \dot{p}))]$ ', the joint denial equivalent of ' $p \vee \sim p$ '.

¹ Wittgenstein calls tautologies 'propositions' at 4.46, though in many other passages he seems to *contrast* the two notions. Let us hence note for the record that tautologies, if we choose to subsume them under 'proposition', constitute a very special case in that they lack "sense" according to the picture theory. However, we need not pursue this issue any further here; it is the distinct identity criteria for 'proposition₁' and 'proposition₂' that are important, not whether we call the expressions involved 'propositions'. If we do not, we can then simply draw a further distinction between 'tautology₁' and 'tautology₂'.

BEGGING THE QUESTION

By DAVID H. SANFORD

JON, the dominant character in Richard Robinson's 'Begging the Question, 1971' (ANALYSIS 31.4), maintains that there is no defect of reasoning called 'begging the question' aside from a certain prohibited move in the Academic game of *elenchus*, a game now seldom played. In 'On Begging the Question at any Time' (ANALYSIS 32.2), Robert Hoffman refutes this contention by giving a condition sufficient for there being just such a defect. I would formulate the condition like this: An argument begs the question if every conjunction of its premisses from which the conclusion follows has as a conjunct a proposition identical with the conclusion. A degenerate case of such question begging is provided by an argument in which the sole premiss is identical with the conclusion. (Hoffman's formulation is different, for he thinks that the fault of begging the question consists in mistaking a certain kind of non-argument for a genuine argument. His view on this matter can be accommodated by inserting the word 'putative' immediately before each occurrence of 'argument', 'premiss' and 'conclusion' in the condition above and in the next three paragraphs.)

Hoffman thinks that his condition is necessary as well as sufficient for the existence of question-begging. This leaves us with several problems. First, he does not explain his notion of propositional identity. I assume that orthographic identity of two sentences in the same argument is a sufficient condition of their expressing the same proposition and that mutual strict implication is a necessary condition. But one wonders if any condition is both necessary and sufficient. If mutual strict implication is regarded as a sufficient condition of propositional identity, then on Hoffman's account too many single premiss arguments are counted as question-begging. If the single premiss is a necessary truth, for example, the argument must be either question-begging or invalid. (See the second page of Robinson's dialogue for a discussion of immediate inference.) If, on the other hand, orthographic identity is regarded as a necessary condition for propositional identity, then for any conclusion whatever one can always give a non-question-begging argument of the form ' $\sim \sim p$, therefore, p '. If one is challenged to defend the premiss, he need only give another argument of the same form, ' $\sim \sim \sim p$, therefore, $\sim \sim p$ '. The tedious process, which strikes us as question-begging, can be continued indefinitely without ever using as a premiss some proposition identical to a conclusion drawn earlier. An adequate definition of propositional identity is required if we are to apply Hoffman's account to all cases, and neither orthographic identity nor mutual strict implication does the trick. But even if such a

definition can be supplied, the forthcoming examples are still troublesome.

Consider the example from the first page of Robinson's dialogue. In at least some circumstances, 'God has all the virtues, therefore He is benevolent' would be commonly regarded as question-begging. Yet its single premiss is not identical with its conclusion. Rather few arguments which are commonly thought to beg the question satisfy Hoffman's account. He leaves us much closer to Jon's position than to that of the critical reader who thinks he often recognizes a question-begging argument when he sees it.

Also, the fault of begging the question is too easily avoided on Hoffman's account. Instead of presenting an argument of the form ' q, p , therefore p ', one can always employ a valid argument the premisses of which are true if and only if the premisses of the first argument are true, but which does not, according to Hoffman, beg the question. Instead of presenting the putative refutation

Hoffman says something.

Indeed, he says something false.

Therefore, he says something false.

I can argue

Hoffman says something.

Either he says something false or he says nothing at all.

Therefore, he says something false.

Some arguments of the disjunctive syllogism form beg the question. Some do not. Begging the question, therefore, is not a purely formal matter. An argument formulated for Smith's benefit, whether by Smith himself or by another, begs the question either if Smith believes one of the premisses only because he already believes the conclusion or if Smith would believe one of the premisses only if he already believed the conclusion. A primary purpose of argument is to increase the degree of reasonable confidence which one has in the truth of the conclusion. Not every argument which fails this purpose begs the question, but every question-begging argument fails this purpose.

The argument 'God has all the virtues, therefore He is benevolent', directed toward one who disbelieves the conclusion, would normally beg the question. Only someone extraordinarily obtuse would accept the premiss, agree that benevolence is a virtue, and deny the conclusion. Someone less obtuse might accept the premiss but neither believe nor disbelieve the conclusion. (If asked out of the blue whether God has all virtues, he would say yes. If asked out of the blue whether God is benevolent, he would be unsure.) In such a case, the argument is perfectly in order. In other circumstances, even though the same

argument is directed toward someone who does not disbelieve the conclusion, it would beg the question. If one is unsure whether God is benevolent, and would believe that God has all the virtues only if he believed that God is benevolent, then the argument cannot increase the degree of reasonable confidence he has in the truth of the conclusion. An argument can also beg the question when it is directed toward someone (for example, 'oneself') who already believes the conclusion. The degree of reasonable confidence one has in the conclusion cannot be increased by an argument if one believes a premiss only because one already believes the conclusion.

'All men are mortal, therefore no immortals are men', the immediate inference mentioned by Robinson, would beg the question in some circumstances but not in all. A valid argument from a single premiss to a conclusion which strictly implies the premiss can sometimes increase the degree of one's reasonable confidence that the conclusion is true. In such a case, it does not beg the question. It does beg the question in case the person to whom the argument is directed would not accept the premiss without previously accepting the conclusion.

The charge that an argument begs the question must thus usually be understood with reference to the beliefs of the person, or the sort of person, to whom the argument is directed. Putative arguments of the sort Hoffman characterizes, however, constitute a limiting case. They always beg the question. There are no circumstances in which they might increase the degree of reasonable confidence one has in the truth of the putative conclusion.

The historical discussion in Robinson's dialogue should discourage groundless etymological speculation about the phrase 'begging the question'. It is an odd phrase; but it has had its present sense in English since at least the late seventeenth century, and no proposed substitute is likely to be adopted. I believe my account fits the usual understanding of the phrase. I have given a condition sufficient for the existence of question-begging which is much broader than Hoffman's sufficient condition. I do not claim to have given a condition which is both sufficient and necessary, for I think it beside the point of the current discussion to dwell on the peculiar situations in which, because the conclusion says something about validity, one would accept the argument as valid only if he would accept the conclusion as true. I claim no originality for the above account. For its sources, and for treatments of the topic just mentioned, see W. E. Johnson's discussion of the epistemic conditions of inference in *Logic*, Part II, L. Susan Stebbing's *A Modern Introduction to Logic*, and Max Black's 'Induction and Experience' in *Experience and Theory*, edited by Lawrence Foster and J. W. Swanson, The University of Massachusetts Press, 1970.

URMSON ON RUSSELL'S INCOMPLETE SYMBOLS

By R. K. PERKINS, JR.

J. O. URMSON, in his *Philosophical Analysis*, accuses Bertrand Russell of holding that 'to show that "X" is an incomplete symbol is tantamount to showing that there are no Xs' [7, 30]. I believe that Urmson is wrong. In what follows I wish to show why I think so, and to indicate the probable source of his mistake.

Urmson appears to draw his evidence solely from a passage in *Principia* [6, 71-2] where Russell is discussing class-symbols.¹ He cites that part of the passage which he finds especially 'puzzling' as follows:

... In the case of descriptions it was possible to *prove* that they are incomplete symbols. In the case of classes, we do not know of any definite proof. ... It is not necessary for our purposes, however, to assert dogmatically that there are no such things as classes. It is only necessary for us to show that the incomplete symbols which we introduce as representatives of classes yield all the propositions for the sake of which classes might be thought essential [7, 29].

It is perhaps understandable how one could be led to Urmson's conclusion on the basis of this passage: one might take the second and third sentences as meaning that *if* there were a 'definite proof' that class-symbols are incomplete, then one *could* 'assert dogmatically' that there are no classes. And, indeed, if this is Russell's meaning, then it surely does look as though he thought that 'to show that "X" is an incomplete symbol is tantamount to showing that there are no Xs'. But I think that it is of crucial importance to understand what Russell has in mind by 'definite proof' here; specifically, how does one go about getting a definite proof that a symbol is incomplete? For Russell there seem to be at least two ways: (A) by showing, as he does in the case of definite descriptions, that the symbol has no meaning in isolation, i.e. that it is not a proper name in his strict and special sense;² or (B) by showing that the object which the symbol putatively symbolizes is non-existent; for if an object is non-existent it cannot, for Russell, be genuinely named, and so its symbol would have to be incomplete. Now the first way, as Russell surely knew, does not, *per se*, permit one to assert that there is no object corresponding to the incomplete symbol. To suppose that he thought so is to suppose that he thought that *all* objects describable by definite descriptions (including, of course, sense-data)—and not just fictitious ones like the golden mountain and the King of France—are

¹ In his excellent book on Russell's early philosophy [1, 24-5] D. F. Pears has a note in which he denies Urmson's general accusation, although he maintains, wrongly I contend, that in the *Principia* passage from which Urmson draws his evidence, Russell makes a 'slip'.

² I have argued elsewhere [2, 45-51] that Russell's argument in *Principia* (p. 67) that descriptions 'mean nothing' is really an argument that they are not genuine proper names.

non-existent. But the second way is quite another matter. For if a symbol has been proved to be incomplete in *this* way, then it follows *ipso facto* that the object which is putatively symbolized must be non-existent.

Now B is really of some importance in connection with the *Principia* passage as cited by Urmson because, in citing the passage, Urmson does a piece of editing which warrants attention; his hiatus between the second and third sentences leaves out a sentence and a footnote which make it clear that Russell was actually thinking of a proof obtainable in this way. He says, 'In the case of classes, we do not know of any equally definite proof, though arguments of more or less cogency can be elicited from the ancient problem of the One and the Many'. And here he adds the following footnote:

Briefly, these arguments reduce to the following: If there is such an object as a class, it must be in some sense *one* object. Yet it is only of classes that *many* can be predicated. Hence, if we admit classes as objects, we must suppose that the same object can be both one and many, which seems impossible. [6, 72fn.]

Here, then, we have an argument which purports to show that the notion of class is self-contradictory. And the argument, if sound, would provide, along the lines of B, a 'definite proof' that class-symbols are incomplete, since classes, being 'impossible', could not exist, and so, could not be named.¹ And it *does* seem to follow quite generally (*pace* Urmson) that if there were such a definite proof of 'X' 's incompleteness, then one *could* 'dogmatically assert' that there are no Xs. Of course, such proofs would only be as good as the arguments purporting to show that X is self-contradictory; and in the case of classes, doubt about the 'cogency' of the argument amounts to doubt about the assertion that classes do not exist. But as Russell says, certainty about this assertion is not required for his purposes: he can get on with the *Principia* programme without either having to affirm or deny that there are such things. In fact, Russell makes it very clear in several other works that the symbols which he calls 'incomplete' are of objects towards the existence of which his attitude is officially agnostic [e.g. 3, 184; 4, 271-4]. Thus, it is generally false to say, as Urmson does, that for Russell:

... the expression 'incomplete symbol' acquired a meaning ... which included the implication not merely that such a symbol was theoretically

¹ I believe that this suggests, incidentally, why Russell felt that he could not get a definite proof of the incompleteness of class-symbols along the lines of A. A proof of type A (see *Principia*, p. 67) requires in its premisses that there be true identity sentences of the form '—is ...' (e.g. 'Scott is the author of *Waverley*') where the right hand side is to be filled in by the symbol to be proved incomplete, and the left hand side is to be filled in by an appropriate proper name, viz. one which names the very same object as the one designated by the symbol on the right hand side. But owing to arguments like the one in the footnote in *Principia*, p. 72, Russell had serious doubts whether there were such things as classes, and so, whether there were any true identity sentences of the required form.

superfluous but that *there was no basic reality for which it stood* [7, 31] (emphasis added).

I believe that the principal source of Urmson's mistake is the fact that Russell frequently, although not in the *Principia* passage, speaks of classes as 'fictions'. Everyone knows that Russell was inclined to ignore ordinary linguistic usage, and frequently employed familiar expressions in odd and not thoroughly explained senses. This, I think, is the case with his use of 'fiction'. Ordinarily, to call something a fiction is to imply that it is fictitious or non-existent. This, however, is not Russell's meaning. For example, when he calls the objects described by definite descriptions 'fictions', he means merely that their symbols are incomplete. This is made quite clear in a passage about classes in *Introduction to Mathematical Philosophy* where, after stating that he wants a definition of class-symbols 'on the same lines as the definition of descriptions', he says:

We shall then be able to say that the symbols for classes are . . . not representing objects called "classes", and that *classes are in fact, like descriptions, logical fictions, or (as we say) "incomplete symbols"* [3, 182; cf. 4, 253] (emphasis added).

Of course, this passage exhibits Russell's frequent carelessness with use and mention. But this can be easily fixed. What he should have said is that 'classes are in fact, like the objects described by definite descriptions, logical fictions, i.e. their putative names are (as we say) "incomplete symbols"'. In this passage Russell intends to say that objects described by definite descriptions, i.e. *all* such objects, are fictions. But he certainly does not intend to imply thereby that all (or any) are non-existent.¹

Now in fairness to Urmson, there is a sense in which he is right in holding that Russell thought that in the case of incomplete symbols 'there was no reality for which it stood'. For the expression 'stood' is crucially ambiguous and capable of yielding at least two very different meanings for Urmson's sentence, viz. (a) There was no reality to which it *referred*; (b) There was no reality for which it was a *name*. Many incomplete symbols may, on Russell's view, be said to 'refer' (or 'designate' or 'denote') in a perfectly familiar sense of that word. (Indeed, Russell is explicit in allowing descriptions, which he sometimes calls 'denoting phrases', to 'denote' [5, 41].) But in his semantics no incomplete symbol can *name*, because only names can name, and incomplete symbols are not names. Thus, we may, with Urmson, say that as regards an incomplete symbol, Russell held that 'there was no reality for

¹ It is true, as Pears has observed [1, 17f and 110] that Russell's term 'fiction' often (though not in the case of descriptions) conveys a point about the kind of analysis employed in the definition of the corresponding symbol, viz. that the analysis is parsimonious, i.e. that it is intended to purge the symbol of that part of its popular meaning which is "illegitimate" and unnecessary. But Russell is officially agnostic about the existence of anything corresponding to what his parsimonious analyses shave away.

which it stood'. But we must realize that Russell held this because incomplete symbols are not names, and *not*, as Urmson seems to think, because he thought that the objects putatively referred to by those symbols are non-existent.

REFERENCES

- [1] Pears, D. F. *Bertrand Russell and the British Tradition in Philosophy* (Fontana Library, 1967).
- [2] Perkins, R. K. 'On Russell's Alleged Confusion of Sense and Reference', *ANALYSIS*, 32.2 (December, 1971).
- [3] Russell, B. *Introduction to Mathematical Philosophy* (London: Allen & Unwin, 1919).
- [4] Id. 'The Philosophy of Logical Atomism', reprinted in R. C. Marsh (ed.) *Logic and Knowledge* (New York: Macmillan, 1956), 178-281.
- [5] Id. 'On Denoting', reprinted in *Logic and Knowledge*, 41-56.
- [6] Whitehead, A. N. and Russell, B. *Principia Mathematica* (Paperback Edition to *56; Cambridge: University Press, 1962).
- [7] Urmson, J. O. *Philosophical Analysis*: (Oxford University Press, 1956).

Duke University

UTILITARIANISM, OVERALL OBLIGATORINESS AND DEONTIC LOGIC

By R. EUGENE BALES

IN a number of recent articles,¹ it has been argued that certain principles of deontic logic reveal that some features of act-utilitarianism are, or may be, problematical. In 'A Problem for Utilitarianism' (*ANALYSIS*, vol. 28 (1968), pp. 141-142), for example, Hector-Neri Castañeda argues that one of the principles of deontic logic, *viz.*, that obligatoriness distributes through conjunction (see (DC) below), and a utilitarian principle such as (U) (also below), together yield contradictory results:

(DC) 'X is obliged_i to do both A and_c B' entails 'X is obliged_i to do A and_c X is obliged_i to do B'.

¹In addition to the articles cited in the text, see Dag Prawitz, 'A discussion note on Utilitarianism', *Theoria*, vol. 34 (1968), and 'The alternatives to an action', *ibid.*, vol. 36 (1970); Lars Bergström, 'Alternatives and Utilitarianism', *ibid.*, vol. 34 (1968); Fred Westphal, in *ANALYSIS* 32.3; H.-N. Castañeda, who provides further bibliographical references, and Harold N. Zellner, both in *ANALYSIS* 32.4.

- (U) X is morally obliged to do A in circumstances C, if and only if X's doing A in C will bring about a greater balance of good (pleasure, etc.) over bad (pain, etc.) than his doing any other alternative action open to him in C.

Castañeda proceeds by asking us to assume that X is morally obliged to do a "conjunctive" act A & B in circumstances C. Then, according to (DC), X is morally obliged to do A and X is morally obliged to do B. But, Castañeda argues, (U) then implies both (a) that the consequences of A are better than the consequences of any alternative (and B in particular), and (b) that the consequences of B are better than the consequences of any alternative (and A in particular). Because of these contradictory results, Castañeda concludes that utilitarian principles cannot provide necessary conditions for obligatoriness.

Some responses to Castañeda's article have taken the form of suggestions that the problem can be circumvented if careful enough attention is paid to the notion of *alternatives* (see e.g. Lars Bergström's 'Utilitarianism and Deontic Logic', ANALYSIS, vol. 29 (1968), pp. 43-44). As far as I know, however, no commentator has pointed out explicitly that the problem arises out of no uniquely utilitarian features of utilitarianism and therefore should not be considered peculiar to utilitarianism.

It should be mentioned that Castañeda does not claim that the problem is one for utilitarianism *alone*. My express claim, however, is that the problem is *not* confined to utilitarianism.

The non-utilitarian theory I have in mind, which shares with act-utilitarianism the problem Castañeda emphasizes, is a deontological theory sometimes suggested as an elaboration of Ross' theory of *prima facie* duties. Indeed, the basis of the theory is a passage from Ross:

- (I) It seems clear that there are cases in which some other *prima facie* duty overrides the *prima facie* duty of fulfilling a promise . . . (T)he ground of the actual rightness of the act is that, of all acts possible for the agent in the circumstances, it is that whose *prima facie* rightness in the respects in which it is *prima facie* right most outweighs its *prima facie* wrongness in any respects in which it is *prima facie* wrong.¹

Taking our cue from Ross, we formulate the following account of overall rightness:

- (OR) An act A is morally right (overall) for an agent X to do in circumstances C if and only if there is no alternative B open to X in C such that B has a greater balance of *prima facie* rightness over *prima facie* wrongness than does A.

Now the question may arise of what we are to say when, as seems

¹ W. D. Ross, *The Right and the Good*, Clarendon Press, 1930. Relevant passages reprinted in Sellars and Hospers, *Readings in Ethical Theory*, Appleton-Century-Crofts, Inc., New York, 1952, p. 196.

clearly possible (at least theoretically), two or more acts A, B, . . . , N, open to X in circumstances C, have an equal balance of *prima facie* rightness over *prima facie* wrongness, but a greater balance of *prima facie* rightness over *prima facie* wrongness than any other alternatives P, Q, One plausible response to the question is that, when such a state of affairs obtains, X's *overall* obligation is to perform one of the acts A, B, . . . , N, but that which particular one of the acts he performs is a matter of moral indifference; in cases where only one right act is open to him, he is obliged (overall) to perform *that* act. Essential features of this response are captured by the following principle:

- (OO) X is morally obliged (overall) to do A in circumstances C if and only if A has a greater balance of *prima facie* rightness over *prima facie* wrongness than does any other alternative action open to him in C.

We now follow Castañeda's line of argument *mutatis mutandis* to show that (DC) and the non-utilitarian principle (OO) together yield contradictory results. We assume that X is morally obliged (overall) to do a "conjunctive" act A & B in circumstances C. Then, according to (DC), X is morally obliged (overall) to do A and X is morally obliged (overall) to do B. But (OO) implies both (a) that A has greater a balance of *prima facie* rightness over *prima facie* wrongness than does any alternative (and B in particular), and (b) that B has a greater balance of *prima facie* rightness over *prima facie* wrongness than does any alternative (and A in particular). Because these results are contradictory, we should conclude—if we are to follow Castañeda's line of argument—that the deontological theory outlined above cannot provide necessary conditions for obligatoriness.

Probably utilitarians have been more persistent than non-utilitarians in their attempts to provide an account of moral obligatoriness in terms of the right alternative (or alternatives) open to an agent in given circumstances. They may take comfort, however, in the observation that utilitarianism fares no worse than does a deontological theory which proposes such an account of obligatoriness. And Bergström (and others) may find satisfaction in the observation that their work on the notion of *alternatives* has application not only within the context of teleological ethics.

Menlo College

MISS HAIGHT ON PICTURING UNREALITY

By W. CHARLTON

IN the way in which an artist can represent the subject of his picture as snubnosed he cannot represent him (i) as real or fictitious (ii) as a particular individual. Miss M. R. Haight (*ANALYSIS*, Jan., 1972) names me as champion of these theses, and offers counter-instances. I reply to her challenge.

Against (i) she offers a picture in which a small girl is leading a large bird which exists only in her imagination. This is no more a counter-instance than a picture showing, *inter alia*, a face in a mirror, a bust of Shakespeare, or a decoy duck. According to (i)—loosely formulated, I dare say, in my *Aesthetics*—an artist cannot represent the depicted scene as a whole as real or imaginary, neither can he represent things in it as real or imaginary in relation to us. In Miss Haight's picture only the bird and the lead are represented as imaginary, and they are represented as imaginary in relation to other things in the picture, or, as Miss Haight herself puts it, 'in a world where the pavement and the child are real'. This world is not represented either as our own real world or as an imaginary one; neither, then, is the girl's exercise of imagination represented as a real, historical one or as a fictitious or imaginary one.

Miss Haight speaks of conventions, and there might be a convention according to which scenes painted in faint, languid colours (cf. Hume, *Treatise* I. iii. 7) are to be taken not as foggy but as fictitious, and scenes painted in bright lively colours not as sunlit but as historical or real. In the first place, however, the convention would be not pictorial but grammatical: using pallid colours, like using 'It is not the case that' would be simply denying, and result not in pictures but in sentences. And second, the scene represented in pallid or bright colours could not be represented as being represented in pale or bright colours.

Against (ii) Miss Haight offers a picture of Lady Batrose as Diana and a picture of Mr. Heath as Britannia; I do not think she will object to my substituting for the latter Scarfe's picture of Mr. Macmillan as Miss Keeler. Again these are not genuine counter-instances. Lady Batrose is not represented as being the same person as Diana in the way in which Christ has been thought to be the same person as the second person of the Trinity. Scarfe does not invite us to think that Mr. Macmillan and Miss Keeler were one and the same person in the way in which the Friar and the Duke in *Measure for Measure* were one and the same person. Miss Haight is right that to represent Mr. Macmillan as Miss Keeler is not simply to represent him as having some of her attributes, but neither is it to represent him as having something more in common with her. It is to represent him as having some of her attributes *in a certain context*

and for certain purposes: with the intention that the beholder should think of them together, and perhaps try to adapt, so as to fit Mr. Macmillan's administration, things currently being said about Miss Keeler. This last remark is itself an adaptation of what has been said about metaphor,¹ and the pictures Miss Haight has in mind may be compared with speeches in which a particular individual is addressed as Venus or Jezebel, or referred to as a wild beast or flower or the like. Such speeches are not identity claims.

The weakness of Miss Haight's counter-instances to (ii) may be seen if we reflect that there are pictures which really might prompt us to say: 'That suggests that A was the same individual as B', 'According to that picture A and B were identical'. So, for instance, a picture of Bacon showing him adding to a pile of manuscript on which the title 'Hamlet' is visible, while behind him hang a mask with a high forehead and other articles of disguise. But even this picture would not be a counter-instance. We might take it as evidence that Bacon was the same person as Shakespeare, but it does not represent Bacon as being the same person as Shakespeare. We should not even say that it represents a man as (a) writing *Hamlet* and (b) being Bacon. Rather it represents a man who is, i.e. who is supposed to be, Bacon, as doing something often thought to have been done by someone else. The identity of Bacon with the author of *Hamlet* may appear from but does not appear in the picture.

¹ M. Black: 'Metaphor', *Aristotelian Society Proceedings*, 1954-5.

University of Newcastle upon Tyne

CENSORSHIP AND PARTICIPATORY DEMOCRACY: A PARADOX

By W. A. McMULLEN

AN *extreme participatory democracy* I define as an organization in which all (adult) members participate in all the decision-making of the organization; and an *informed* extreme participatory democracy I define as one in which all the members have access to all the information relevant to all the decisions that the organization has to take. This concept of an informed extreme participatory democracy is, I think, an ideal in the minds of many people today; an ideal which they hold because of the large size of most present-day democracies and the consequent distance of the average individual from most decision-making and from much of the information on which many decisions are based.

Such an extreme form of democracy, however, runs into a problem concerning censorship. Take pornography as an example. If all the citizens are to participate fully in all decisions, they will all have to scrutinize thoroughly all the pornographic matter that they are considering banning, and they will thus be engaged in a self-defeating enterprise. Censorship cannot work for them, since by the time they have performed their work as censors everyone will have seen all the doubtful material; and, if it is indeed harmful, will already have been harmed by it. Of course, they can still make decisions as to what should be kept away from children; but censorship for the protection of adults is ruled out by the logic of the concept of informed extreme participatory democracy as defined above.

More accurately, no pornographic material could be kept *totally* away from any adult in such an extreme democracy. It *could* be kept *partly* away from some or all adults, since, after the scrutinizing process was over, the organization could decide that some or all of the offending material was so odious that they would ban it henceforth (either for all their number or for those presumed likely to be harmed further by it). This possibility means that censorship in such a democracy would not be completely self-defeating. For in such a case what the organization would be doing would be judging that, although harm had been done by the scrutinizing process, yet more harm would ensue if the pornography remained available.

An informed extreme participatory democracy also runs into a difficulty over state secrets, if it has any. In this case all the citizens could indeed know all of the state's military secrets without contradiction. For here they are only concerned with keeping material from potential enemies, not from themselves. However, such an extreme democracy needs loyalty from every one of its citizens: one traitor could sell every one of its secrets. This is ironical: universal loyalty is usually thought of as a demand of totalitarianism rather than a need of participatory democracy.

These difficulties regarding censorship and state secrets show that democracies with less than complete participation in decision-making do have some theoretical advantages vis-à-vis extreme forms of participatory democracy. Indeed, though the following consideration will at some point be outweighed by others, it will certainly up to a point be true that, if obscene material ought to be banned, then the fewer citizens involved in the process of censorship the better. And, again, although this principle too will at some point be outweighed by others, up to a point the fewer people with access to state secrets the better.

NOTES

The ANALYSIS Committee consists of: Chairman, P. T. Geach; Secretary, R. G. Swinburne; J. H. Benson, Miss M. A. Boden, Alec Fisher, Andrew Harrison, R. F. Holland, Stephan Körner, A. C. MacIntyre, A. M. MacIver, Bernard Mayo, A. R. White, Peter Winch. This committee is responsible for appointing and advising the editor and for the general policy of the paper.

SUBSCRIPTIONS. The subscription to ANALYSIS, Vol. 32, is £1.50 for individuals in Great Britain, £1.75 for individuals abroad, and £2.00 for all institutions, post free. Each volume comprises six numbers, five of 32 pages and one of 48 pages, appearing within the academic year—in October, December, January, March, April and June. Orders should be sent to Basil Blackwell, 49 Broad Street, Oxford, or placed with any bookseller.

CONTRIBUTIONS. Articles submitted for publication should be addressed to Dr. C. J. F. Williams, 1 Fossefield Road, Midsomer Norton, Bath, BA3 4AS. Contributors are asked to note the following requirements.

Articles should **normally** not exceed 3,000 words in length. **Occasionally**, however, longer contributions can be accepted;

They must be typewritten in double spacing on one side of the paper only;

Footnotes should be kept to a minimum and wherever possible avoided altogether;

Single quotation marks should normally be used, except for purposes of internal quotation and "scare" quotes.

Discussion papers should be sent in as soon as possible after the appearance of the article to which they refer.

Contributors in the United Kingdom should enclose a stamped addressed envelope of suitable size; if immediate acknowledgement is required a stamped postcard should also be enclosed.

Overseas contributors who wish to have their MSS. returned should send an envelope **and international reply coupons** of the requisite value, whether for air or surface mail.

Proofs are normally read by the Editor. A contributor who wishes to read his own proofs should notify the Editor at the same time as the MS. is submitted.

OFFPRINTS. Contributors will be supplied free of charge with a limited number of offprints of their articles. Further copies may be ordered at a charge from the publishers. Orders for these, if required, should be placed **direct with the publishers** when an article has been accepted for publication.

The copyright of articles printed in ANALYSIS remains the property of the author, but contributors are strongly advised, in their own interest, to consult the Editor before consenting to the reprinting of their articles.